Financial Instruments and Risk Management

2013

P. Ouwehand

Department of Mathematical Sciences
Stellenbosch University

Contents

Chapter 1

Basic Concepts in Finance

In this chapter we introduce some of the basic theory, terminology and conceptual machinery of finance. As the ideas introduced here will be important for many of your courses, it is important to rapidly master them.

Finance may be defined as *the study of how people allocate scarce resources over time.* Individuals and corporations may use savings accounts, mortgages, pension funds, annuities, stocks, etc. for investment purposes. The outcomes of financial decisions — the *costs* and *benefits* — are usually:

- spread over *time*;
- not generally known with certainty ahead of time, i.e. subject to an element of *risk*.

To make intelligent investment and consumption decisions, individuals must be able to *value* and compare different *risky* cashflows over *time*.

Essentially finance is rests on three pillars:

(1) Time value of money

(2) Valuation (of assets, stocks, bonds, derivatives)

(3) Risk management

You will meet aspects of each of these in this course.

1.1 The Time–Value of Money

Because financial decisions involve costs/benefits that are spread over time, decision makers must be able to compare the values of different projected cashflows at different dates.

Example 1.1.1 Which do you prefer?

A. \$1 000 in hand today, or

B. The promise of \$1 000 in one year's time?

Why?

□

Most people would prefer option A of the above example. There are at least three reasons for this:

- *Opportunity cost:* You can *invest* the \$1 000 now, with the *expectation* of receiving a greater sum in the future. (Time is money.)
- *Inflation:* The purchasing power of money changes (generally decreases) over time. \$1 000 in one year's time may buy fewer goods than \$1 000 today.
- *Risk/Uncertainty:* You cannot be absolutely sure that you will actually receive the \$1 000 in one year's time.

The "time value of money" refers to the fact that a fixed sum of money in the hand today is worth more than the *expectation* of receiving the same amount at some future date. Basically, a person who lends money is giving up the opportunity to convert those funds into consumable goods or services. So borrowing isn't free: The borrower must pay a premium to induce the lender to part temporarily with his/her money — the *interest*. The *interest rate* depends on many factors, e.g. inflation, money supply, credit rating, etc.

1.1.1 Interest

The magnitude of interest rates depends on:

- Economic factors (inflation, trade deficit figures, etc.)
- Credit rating of borrower: The interest charged is higher if the borrower is perceived to have a greater chance of default. This is because investors require an additional premium for the risk they take on. Thus government debt (Treasuries/Gilts) pay lowest rates ("risk-free" investments), whereas "junk"– or non–investment grade bonds pay the highest.

Junk bonds are issued by corporations with a high risk of going bankrupt, i.e. defaulting on their payments. The fact that investors receive higher interest from corporate debt than from government debt illustrates one of the most fundamental ideas in finance: There is a direct relationship between the *risk* of an investment and its *expected* return. The higher the perceived risk, the greater the risk premium that investors expect in order to induce them to part with their money. Governments have almost no chance of defaulting on the bonds they issue (in the local currency, that is): After all they can always raise more money by raising taxes, or even by simply printing more money[1]. For that reason, these are frequently called *risk–free* investments — a misnomer, because although there is no chance of default, they do carry some other forms of risk: primarily interest rate risk (the values of bonds may fluctuate wildly), liquidity risk (for some bonds with very long maturities) and FX risk (for foreign investors).

Interest rates vary over time, and vary for different loans at the same time, but for the moment we will assume that they are constant. Suppose that the annualized interest rate is r. We distinguish between three kinds of interest:

[1] This is not always possible, though. The Zimbabwean dollar/ US dollar exchange rate plummeted from 650 Z\$ per US\$ in July 2006 to 750 *billion* Z\$ per US\$ in July 2008. The main reason was the the Zimbabwean government was printing enormous amounts of money. Printing money was done to such a great extent that the printing presses were constantly breaking down and that they ran out of paper for printing bank notes.

(a) **Simple interest:** If a sum P_0 is deposited at a (simple) rate of r p.a., it will grow to $P_0(1 + rT)$ after T years.

(b) **Discretely compounded interest:** If the interest is compounded n times per year, you will have $(1 + \frac{r}{n})^n$ after one year, and $(1 + \frac{r}{n})^{nN}$ after N years.

(c) **Continuously compounded interest:** Suppose that the interest is compounded at increasingly frequent times. Deposit \$1.00 for 1 year. As the number of times it is compounded, n, increases, we get $\lim_{n\to\infty}(1 + \frac{r}{n})^n = e^r$ after 1 year, and e^{rT} after T years.

We will use all of these types of interest rate in this course. It is mathematically convenient to work with continuously compounded interest rates, for example. However, the rates quoted in the money market (LIBOR, JIBAR) are generally simple. Most bonds pay a semi–annual coupon, and thus the semi–annually compounded rate is the appropriate one to use when valuing bonds.

Two related concepts that we will use are present value and future value. The *future value* of a sum P invested for N years at an interest rate r compounded annually is simply

$$FV = P(1 + r)^N$$

The *present value* of a sum P is simply the sum we have to invest now to get P at a future date. For example, the present value of \$1000.00 to be received in 5 years time at an (annually) compounded interest rate of 10% is:

$$PV = P(1 + r)^{-N} = \frac{1000}{1.1^5} = 620.92$$

Present value and future value provide the means by which cashflows can be compared over time.

$$FV = PV(1 + r)^T \qquad \text{or} \qquad FV = PVe^{rT}$$

Example 1.1.2 Suppose that we are valuing an investment that will pay \$1000.00 each year for the next 10 years, starting one year from now. If a bank would pay us 8% p.a. on a 10–year deposit, then the present value of this investment is

$$PV = \sum_{n=1}^{10} \frac{1000}{(1.08)^n} = 1000\frac{1 - \frac{1}{1.08^{10}}}{0.08} = 6\ \ 710.08$$

Thus you should not pay more than \$6 710.08 for this investment.

□

Calculating present values is known as *discounting*, and the interest rate used in the calculation is often called the *discount rate*. Calculating future values is known as *compounding*.

Because the frequency of compounding can differ, it is important to have a way of making interest rates comparable. One must therefore be able to convert an interest rate from one compounding frequency to another. In practice, this involves a trivial calculation.

Example 1.1.3 • *What is the present value of an investment that will pay \$50 semi–annually for the next 3 years if the semi–annually compounded rate for all maturities is 12%?*

$$PV = \sum_{n=1}^{6} \frac{50}{(1.06)^n} = \$245.87$$

- *Which investment is preferable: One offering 10% compounded quarterly, or one offering 9.9%, compounded continuously?*
  To find the continuously compounded rate, r, which corresponds to a quarterly rate of 10%, we must solve
  $$e^r = (1 + \frac{0.10}{4})^4$$
  which yields $r = 9.88\%$. Hence the investment offering a continuously compounded rate of 9.9% is a little better.

□

1.1.2 Returns

In this this section, we briefly discuss the notion of the *returns* of an asset. The rate of return of a risky asset is mathematically very similar to an interest rate. The main difference is that the interest rate is a *promised* return on a deposit, whereas the returns on an asset do not involve a such a contract, and therefore carry more risk. When you invest in equity (shares), the return comes from two sources: The cash *dividend* that is paid out to the shareholders by the corporation that issued the stock, and the gain (or loss!) in the market price of the stock over the period that it is held[2].

Suppose for example that you bought a share of Xcor one year ago for R123.45. Today the stock pays a dividend of R12.00, and the share price is R135.40. The total gain is therefore R$(135.40 - 123.45 + 12.00) =$ R23.95. The *rate of return* is $\frac{23.95}{123.45} = 19.4\%$.

Now compare this to the following calculation: One year ago you deposited R123.45 in the bank. Today you withdraw R12.00 and there is still R135.40 left in the bank account. A quick calculation will show that the corresponding simple interest rate must be 19.4% as well.

Basically, in the above situations, the simple rate r of return is given by

$$r = \frac{\text{Final Share Price} - \text{Initial Share Price} + \text{Dividends}}{\text{Initial Share Price}}$$

Naturally, if the share is held over an extended period, with dividends paid out at various times, one must perform slightly more complicated calculations.

One can therefore think of the return on an asset as the interest earned on that asset. One must also differentiate between *expected* and *historical* returns. The *expected return* of a share is the average (but uncertain) return expected in the future by investors, whereas the historical return is the return that was actually achieved by the share. The latter may well be negative! The former cannot be negative, because nobody would buy the share. Generally, the riskier the perceived future returns of a share, the higher its expected rate of return: Investors require a risk premium in return for taking on the risk of buying the share. For

[2]The gain(loss) due to the change in market prices is called *capital gain(loss).*

that reason the average (expected) rate of return on a share is higher than the interest rate on a bank deposit.

Now in the above case, the rate of return corresponds to a simple interest rate. Do we also have notions of return that correspond to copound rates? We do indeed. Suppose that we pay an amount of \$400 today, and in return will receive \$100 at the end of each year for the next 5 years. The *internal rate of return* is defined to be that annualized rate r which makes the present value of the future cashflows equal to the cost. Thus r is the solution of the equation

$$400 = \sum_{t=1}^{5} \frac{400}{(1+r)^t}$$

To find r, we therefore have to find the roots of a quintic, i.e. a fifth degree polynomial. We therefore see that

- There may be more than one internal rate of return, because the polynomial may have more than one root.
- We will have to use a numerical method to calculate "the" internal rate of return: The Norwegian mathematician Niels Henrik Abel proved in 1824 that there is no formula (involving radicals) for the solution of a general quintic. Fortunately Excel has a number of useful numerical routines, including the function IRR, which tells us that $r = 7.93\%$.

Another way in which returns are sometimes defined is as follows (ignoring dividends):

$$\text{Rate of Return } R = \frac{1}{T} \ln \frac{P_T}{P_0}$$

where P_T, P_0 are the prices at times $t = T, t = 0$ respectively. These *log-returns* correspond to looking at returns as a kind of continuously compounded interest, because $P_T = P_0 e^{RT}$.

1.2 Financial Markets and Instruments

Traders in a financial market exchange securities for money. (*Securities* are contracts for future delivery of goods or money, e.g. shares, bonds and derivatives). The main players in a financial market are

- Individuals or households
- Corporations
- Financial intermediaries, e.g. banks, investment companies (mutual funds, pension funds), insurance companies and stock exchanges.

Individuals and corporations each face some of the following financial decisions:

(a) Consumption/savings decisions: How much of current wealth should be consumed, and how much saved for future consumption.

(b) Investment decisions: What to do with savings.

(c) Financing decisions: How to use other people's money to achieve aims. Corporations often must decide between how much money they will raise by taking loans, and how much by issuing shares.

(d) Risk management decisions: How to reduce future financial uncertainties.

The financial system (comprised of markets and intermediaries) is the means through which these financial decisions are carried out.

Apart from *buyers* and *sellers*, there are a number of different roles that may be played by agents in financial markets:

- *Market makers* are firms or individuals that *make* markets by linking up buyers and sellers and coordinating trades. Market makers quote
  - a *bid price*, at which they are willing to buy a security, and
  - an *ask price* (or *offer price*) at which they are willing to sell a security.

  Market makers are obliged to buy and sell at their quoted prices. Market makers do not typically hold inventories of a large number of securities or trade for their own account. The revenues they generate are derived mainly from the *bid–ask* spread — the difference between the higher ask price at which they sell and the lower bid price at which they buy.

  A market maker may make a market in just one security, or in a number of different securities.

- *Traders* buy and sell securities on behalf of clients.

- *Brokers* will also act as intermediaries for buyers and sellers, and execute trades. They will also offer other services, such as advice on investments. They do not hold inventories of a large number of securities.

Some trades are done on organized *exchanges* such as stock exchanges or futures exchanges. An exchange is a formal entity where buying and selling takes place. Trading products and methods are *standardized.* This reduces transaction costs, but may also reduce flexibility in tailoring trades to meet specific needs. An exchange may be *electronic* or *open–outcry.* In some cases, the clearing house of the exchange becomes the *de facto* counterparty to each trade — thereby significantly reducing credit risk.

Other trades are done *over–the–counter* (OTC). OTC markets are not formally organized (though they are regulated), and the particulars of each deal can be carefully negotiated between counterparties.

- One usually distinguishes between *primary* and *secondary* markets: Securities sold (issued) for the first time are sold on the primary market. If those securities can then be traded again, this is done on the secondary market. For example, in the USA the government borrows money by issuing Treasury bonds. These bonds are, on issue, auctioned to a select group of institutional investors — the primary market. Afterwards, these bonds can be bought and sold freely by any one on the secondary market. If there is an active secondary market, an investor who entered a contract in the primary market is not locked in, but can exit this contract by selling in the secondary market. Thus the secondary market provides liquidity, and is the main place where investor sentiment makes itself felt. By doing so, it also influences the primary market.

- One also distinguishes between *underlying* (primary) and *derivative* (secondary) instruments. Examples of underlying instruments are shares, bonds, currencies, interest rates, and indices. Their values are determined by supply and demand. A derivative is a financial instruments whose value is *derived* from an underlying asset (or basket of assets)[3]. An example of a derivative that we will treat in great detail is that of a (European) *call option*. This gives the holder of the option the *right* (but not the *obligation*) to buy the underlying at a fixed price (the strike price) at some predetermined future date (expiry).

- Borrowing and lending is done in *fixed income markets*. The fixed income market is split into the *capital market*, for debt of maturities ≥ 1 year, and the *money market*, for short–term debt.

- *Euromarkets*[4] — eurocurrency and eurobond markets — are markets where banks deal in currencies other than that of the country in which they are located. For example, an Australian company might issue a dollar–denominated bond in Japan. By using euromarkets, banks are able to avoid certain costs and restrictions imposed by government regulation.

- Finally, we distinguish between the *spot market* and the *forward market*. Most transactions are *spot transactions*, in that payment and delivery of goods occur immediately after the transaction is closed. To hedge against unfavourable moves of the market in the future, it is possible to sell goods now for delivery in the future. For example, a farmer and a bakery can agree today that the farmer will sell 10 metric tons of wheat to the bakery in 3 months' time, for \$3 500. In 6 months' time, wheat prices might be \$400 per ton. Nevertheless, the farmer is obliged to sell at \$3 500 per ton, as agreed.

1.2.1 Equity

One of the most basic class of financial instruments is *equity* (stocks or shares). A share is ownership of a small piece of a company. Anyone who wants to start or expand a company can raise the necessary capital by issuing shares. Investors in the stock (the shareholders) pay cash in return for a share in the profits of the company. These profits are paid in the form of *dividends* and growth in the stock's value. Dividends are lump sum payments, paid out to investors at regular intervals (typically semi–annually). The amount actually paid out varies, depending on the actual profit made by the company and the opportunities for growth. Some fast–growth companies initially pay no dividends, because all the profits are ploughed back into expanding the company. However, in all cases, the price of a share reflects the profit investors eventually expect to reap.

The relationship between risk and return mentioned in the previous section also holds for shares: Shares with a higher expected return are *riskier* than those with a lower expected return. This is essentially because most investors are *risk averse*, i.e. they need to be compen-

[3]These days, though, the name is applied to almost any traded hedging(or speculating) device. For example, weather derivatives allow one to hedge against (or place a bet on) the economic effects of the weather. Though the weather can hardly be seen as an asset in the traditional sense, many businesses do profit from it (e.g. Hawaii, Aspen).

[4]Note that the term *euro* has nothing to do with the continent of Europe, or its currency, the Euro €.

sated to take on risk.[5] Thus if two shares were perceived to have the same risk, but different expected returns, investors would try to sell the riskier share, and buy the less risky share. There would therefore be a net supply of the risky share, causing its price to drop, and its expected return to increase. Similarly, there would be a net demand for the less risky share, causing its price to rise, and its expected return to decrease.

The shareholders are the *owners* of the company. The *directors* of a company (are supposed to) act in the shareholders' best interests. This simply means that they must make as much profit as possible, since the profits will eventually be paid out to the shareholders. Since the share price reflects expected profits, directors try to ensure that the share price is as high as possible. *Public limited companies* are listed on a stock exchange. This facilitates the transfer of ownership (buying or selling shares). Though shareholders of a public limited company are entitled to the profits of a company, they have *limited liability*. This means that they cannot be held responsible for any debt that a company may have if it goes bankrupt.

Shares may be bought *cum*– or *ex*–dividend: If the share is bought before the ex–dividend date, the new owner will be paid the next dividend. Otherwise, the dividend will be paid to the old owner. Thus on the ex–dividend date, the share price drops by aproximately the present value of the announced dividend.

Occasionally acompany will announce a *stock split*. This does not affect the value of the company. Suppose, for example, that you own a single share whose current value is $600.00. After a 3–for–1 split, you will own 3 shares, each valued at about $200.00.

1.2.2 Short Sales

Many financial securities can be *sold short*. This allows investors to sell securities that they do not own, hoping to buy them more cheaply later on. The mechanics of a short sale of shares, for example, are as follows:

- Your broker borrows the share from another client.
- You may now sell this share, even though you do not own it.
- Later, you buy the share in the market and return it to your broker, who returns it to the original owner. You must also pay any dividends that were issued in the interim.

Mathematically, this is significant because it means that an investor can own a negative number of shares.

1.2.3 Commodities

These are products such as foods, metals, oil etc. Their prices are determined by supply and demand. Commodities are traded not only by companies that need them for production, but also by people who speculate on the direction of price movements.

1.2.4 Fixed Income Securities

The term *fixed income security* originally referred to fixed coupon bonds, but is now used to cover many interest rate related securities.

[5] Consider the following bet: A fair coin is flipped. If it lands heads, you get $1 million. If it lands tails, you have to pay $1 million. Most people would not be prepared to take the bet, even though it is fair. They would have to be paid something in order for them to agree to take this bet. This "something" is the *risk premium*.

Bonds

Bonds are an example of *fixed income securities*, and are essentially just loans (to the government or to a company). They promise to pay, in return for an initial deposit, a predetermined amount at a predetermined date. They often also pay out predetermined amounts at intermediate times (the *coupon*). For example, a 10–year bond with face value \$100 000 and a 10% coupon, paid semi–annually, will result in a cashflow of \$5 000 at 6 month intervals for times 0.5, 1.0, 1.5,. . . , 9.5, and a final payment of \$105 000 at time 10 years. A *zero–coupon bond* pays no coupons, just a single bullet at maturity. At the other end of the spectrum, a *consol* pays only coupons, and never its face value. It may be regarded as a coupon bearing bond that never matures.

Annuities

An annuity is another example of a fixed–income security. An annuity will pay out a fixed amount at regular intervals in return for a lump sum upfront. For example: You are just about to retire and are considering whether or not to buy an annuity from an insurance company. If you pay \$100 000, the company will pay you out \$10 000 per year for the next 15 years. Whether or not this annuity is worth investing in depends on current interest rates.

Mortgages are another example of an annuity. Here the homeowner pays a fixed amount at regular (e.g. monthly) intervals for a prescribed period (e.g. 20 years) in return for receiving a single payment upfront (the loan to buy the house).

Chapter 5 is concerned with fixed income securities.

1.2.5 Derivatives

A derivative security, often known as a derivative or contingent claim, is a financial instrument whose value is derived from the values of more basic underlying variables, e.g.

Equity prices	—	e.g. stock options
Commodity prices	—	e.g. porkbelly futures
Stock indexes	—	e.g. index futures
Foreign exchange	—	e.g. forward rate agreements
Interest rates etc.	—	e.g. interest rate swaps
Other derivatives	—	e.g. futures options

One important difference between underlying and derivative instruments is the method of pricing: As derivatives are strongly dependent on the underlying asset, their (fair) prices are a function of the value of the underlying, i.e. their prices are not directly determined by supply and demand. It is this difference which makes financial mathematics, the use of mathematics to value securities, possible.

For example, the price of a call option on a share depends on the share price (and also the risk–free rate, the strike price, the time to expiry and the volatitility of the share). One cannot use financial mathematics to calculate the value of a share, but one can use it to find the value of the call option. This works as follows: One can construct a portfolio consisting of bonds and the underlying asset which has the same cashflows as the option. This portfolio is called a *replicating portfolio*, for obvious reasons. It then follows that the portfolio must have the same value as the derivative. If this were not the case, there would be *arbitrage opportunities*. These are essentially "free lunches" – the possibility of making a profit with no chance of a loss. In this case, if the replicating portfolio is cheaper than the call, then (1) sell

the call; (2) use the proceeds to buy the portfolio; (3) pocket the difference. When the call expires, your obligations on the call are matched by the income received from the replicating portfolio. Of course, if the replicating portfolio is more expensive then the portfolio, buy the call and sell the portfolio. Always buy cheap and sell dear! In either case, the difference between the prices of the replicating portfolio and the option is free money, obtained at no risk. The laws of supply and demand will now force the prices of the portfolio and the option to converge.

Forwards and futures on equity and commodities will be discussed in Chapters 3 and **??**, where we also look at the notion of arbitrage in more detail. Fixed income derivatives are the subjectof Chapter 6. The Black–Scholes price of a call option will be derived in Chapter 7.

Chapter 2

Introduction to Financial Derivatives

A *derivative security* can be defined as a financial contract whose value is *derived* from an underlying variable. This variable may be the value of an asset, such as a stock, a commodity, a bond, foreign currency or even another derivative. It may also depend on more abstract financial observables, e.g. economic indicators, the level of an index, or an interest rate. Newer types of derivatives will pay out if a corporation or nation defaults on its debt obligations (credit derivatives). There are even derivatives depending on the temperature data (electricity/weather derivatives) or on whether or not there is an earthquake (insurance derivatives).

2.1 Reasons for Using Derivative Securities

There are two reasons for using derivatives: *hedging* and *speculation*. These are two sides of the same coin. *Hedgers* attempt to reduce risk, whereas *speculators* take a risky position in the hope that it will reap rewards. We shall give some examples of these activities soon. Apart from hedgers and speculators, there is a third important type of trader, namely the *arbitrageur*, who attempts to make profits by entering off–setting positions in different markets or instruments.

Derivatives are tools for *transferring risk*

The use of derivatives will allow you to lessen or increase your exposure to uncertain events.

Before we can give some examples of these trading activities, we need to introduce some simple derivatives.

2.1.1 Forward–Based Contracts

Forward–based contracts entail *obligations* to exchange assets at a future date, or at a set of future dates. These assets might be financial securities, commodities or pre–defined cashflows. Here is a basic example.

Definition 2.1.1 A *forward contract* is an agreement to buy or sell an asset S (the *underlying*) at a certain date T (*delivery date, maturity*) for a certain price F (*forward price, delivery price*).

- The party who agrees to buy the asset is said to have the *long* position, and the party that agrees to deliver the asset is said to have the *short* position.
- Provided that the forward price F is chosen carefully, this contract initially has no value, i.e. it costs neither party anything to enter into the contract. We shall see how to calculate F in the next section. As time goes by, the value of the contract changes.
- Suppose that the *spot price* of the asset at time T is denoted by $S(T)$. The party with the long position therefore agrees to pay F for what is worth $S(T)$. The *payoff* to the holder of the long position is therefore $S(T) - F$. (This may be positive or negative.) The payoff to the party with the short position is $F - S(T)$.

□

Note that the payoffs cancel out: One party's gain is the other party's loss. Hence forwards are a zerosum game.

A *futures contract* is very similar to a forward contract, and for the moment you can regard forwards and futures as pretty much the same thing. We'll discuss the differences in Chapter 3.

Example 2.1.2 (a) BakeCor, a large bakery knows that it will need 1 000 bushels of wheat in 6 months' time. The current price of wheat is R10.00 per bushel. Ignoring interest rates[1], the managers at BakeCor enter into a forward contract with a farmer to buy 1 000 bushels in 6 months' time at R10.00 per bushel. BakeCor does this because the price of wheat 6 months hence is uncertain, depending on a large number of factors, including the weather. If the weather is bad, the wheat crop will be small, leading to scarcity and thus to higher wheat prices. By entering into a forward contract, BakeCor locks into a predetermined price and eliminates this risk. However, this elimination comes at some cost: If the weather is good, wheat prices will be lower than they are now, but BakeCor can no longer profit from this price drop: Its payoff is negative.

The farmer's position is just the opposite: Good weather means a bumper crop, not just for him, but also for all the other farmers. Wheat prices will drop to induce people to consume more wheat products, but the farmer may still end up being unable to sell a portion of his crop. The forward contract protects against this risk.

In this example, both BakeCor and the farmer are using a derivative as a hedging tool.

(b) Investor X has some reason for expecting a bumper crop of wheat this year (perhaps she believes this because of technical analysis, or astrology, or perhaps she genuinely has superior information). She therefore expects the price of wheat to drop drastically, and takes the short position of a forward contract. She agrees to deliver 1 000 bushels of wheat at R10.00 per bushel in 6 months' time. Investor X does not currently own any wheat; she is speculating. She expects to be able to buy the wheat for R7.00 per bushel in 6 months' time and will deliver this to the holder of the long position, thus making a profit of R3.00 per bushel.

If she is wrong about wheat prices however, and it rises to R15.00 per bushel, she will make a loss of R5.00 per bushel.

[1] i.e. assuming $r = 0$. As we shall see in the next section, if $r > 0$, then the forward price will be > 10.00

□

Example 2.1.3 A US corporation knows that it will need to buy equipment to the value of Eu. 1.5 million[2] from a German manufacturer in 2 months' time. In order to hedge foreign exchange risk, the corporation investigates the possibility of entering a forward contract with an international bank. The bank quotes the following USD/Euro exchange rates:

	Bid	**Offer**
Spot	0.9989	0.9993
30-day forward	0.9973	0.9977
60-day forward	0.9956	0.9960
90-day forward	0.9940	0.9944

Thus the bank is offering Euros at 0.9993 dollars per Euro today, and at 0.9960 dollars per Euro in 2 months' time. Thus these exchange rates are *prices*: A bid price of 0.9989 today means that the bank is willing to buy a Euro for 0.9989 dollars.

The US corporation now has two options: It can wait for two months, and then buy dollars at the then-prevailing spot rate, or it can enter into a forward contract.

If it decides to wait, it is subject to foreign exchange risk: In two months' time, the spot rate might be (say) 0.9910, in which case the corporation is better off waiting. On the other hand, it might also be 1.213, in which case the corporation is off worse. The point is that the spot rate in two months' time is unknown, i.e. it is a random variable. By electing to wait, the corporation is taking on risk of not knowing exactly how much it will have to pay for Euros. But this risk can result in a profit as well as a loss.

On the other hand, by entering a forward contract with the bank, the US company knows that it will be able to buy the required Eu. 1.5 million for $1 494 000 in 2 months' time. Since this number is known today, there is no randomness in what will eventually be paid, i.e. the foreign exchange risk is hedged away. However, it is now impossible to benefit from a favourable move in exchange rates (i.e. dollar appreciation): If the rate had dropped to 0.9910, the corporation would have been better off not entering a forward contract.

At this point you may be wondering how the forward rates are determined. Why is the forward rate not simply te same as the spot rate? In fact, Euros seem to be getting cheaper, according to the table! The answer is: Interest rates.

We will discuss the determination of *forward prices* in detail in a later chapter, but the underlying ideas are quite simple. Suppose that the US corporation is able to borrow/lend at 2% in the USA and at 4% in Germany. Today's USD/Euro exchange rate is 0.9993. Now we can do one of two things with $0.9993

1. We can invest it at 2% in the US for two months, yielding $0.9993e^{0.02\times\frac{60}{365}} = 1.0026$ dollars in two months' time.

2. We can use it to buy 1 Euro, and invest that for two months at 4% in a German bank, yielding $e^{0.04\times\frac{60}{365}} = 1.0066$ Euros.

Intuitively, we expect that these investments should come to the same thing, as we shall explain in a moment, i.e. we expect that 1.0026 dollars will equal 1.0066 Euros in two months' time. Hence 1 Euro is expected[3] to be $\frac{1.0026}{1.0066} = 0.9960$ dollars.

[2]Eu. will stand for Euro

[3]What we *expect* is not necessarily what will happen!

Why do we *expect* these investments to come to the same thing? Ignoring credit risk, if the market expects 1 Euro to cost less than 0.9960 dollars in two months' then Euros will be *cheap*. It will therefore profit traders to *buy* Euros in two months' time. Suppose, for example, that you expect the spot exchange rate to be 0.9940. The following strategy is then expected to yield a profit: Borrow 10 000 Euros today at 4% and buy 9993 dollars (ignoring the bid/offer spread). You will need to pay back 10 066 Euros in two months' time. Invest your dollars in a US bank at 2% to obtain 10026 dollars after two months. With these dollars, you expect to be able to buy 10 026/0.9940 = 10 087 Euros. After you've paid back the 10 066 Euros which you owe, you will have made a profit of 21 Euros.

If a large proportion of traders feel that the spot rate will be lower than 0.9960 in two months' time, they will sell Euros and buy dollars. Thus the laws of supply and demand dictate that the Euro will depreciate in value, and the dollar will appreciate. Hence the spot exchange rate (today's rate) will drop. Of course, this argument is still kind of wishy–washy. Later in this chapter we shall use an arbitrage argument to make the above train of reasoning precise, and to remove all your lingering doubts.

□

Interest rate swaps are by far the most widely traded OTC derivatives:

Definition 2.1.4 An interest rate swap is an agreement between two parties to exchange a series of payments at a fixed rate of interest R for a floating rate of interest L (typically based on LIBOR). Given a *tenor structure* of future dates $T_0 < T_1 < T_2 < \cdots < T_n$, and a notional amount N, parties agree that one side will pay a fixed amount $NR\,\Delta T$ at times T_i (for $i = 1, \dots, n$), in return for an as yet uncertain amount $NL_i\,\Delta T$. Here $\Delta_T := T_i - T_{i-1}$, and L_i is the prevailing floating rate for the period $[T_{i-1}, T_i]$, which is known only at time T_{i-1}.

The fixed rate R is chosen so that the initial value of the swap contract is zero.

□

2.1.2 Option–Based Contracts

Definition 2.1.5 An *option* gives the holder the right, but not the obligation to buy or sell an asset.

(a) A *European call option* gives the holder the right to *buy* an asset S (the *underlying*) for an agreed amount K (the *strike price* or *exercise price*) on a specified future date T (*maturity* or *expiry*).

- The party who undertakes to deliver the asset is called the *writer* of the option.
- Because the holder of an option does not have to exercise the option, the payoff to the holder is never negative. For example, the buyer of a European call would exercise at expiry T if and only if the strike K is less than the spot price of the underlying $S(T)$. In that case, he would make a profit of $S(T) - K$. If the spot price is less than the strike, the holder would discard the option (i.e. he would not exercise: Why pay K if you can pay $S(T) < K$?)
- Thus the payoff to the holder is simply $\max\{S(T) - K, 0\}$.

- Since the payoff is never negative, options are not free. Nobody would write an option unless there was some form of compensation! Thus options, unlike forward contracts, have a cost associated with them. You have to pay the writer of an option a *premium* upfront to enter into the contract. This premium can be thought of as the *price of the risk that spot price of the underlying falls below K at maturity.*

(b) A *European put option* similarly confers the right to *sell* an asset for an agreed amount at a specified future date.

(c) Similarly, an *American call (put) option* confers the right to buy (sell) an asset for an agreed amount. However, this right can be exercised at any time before maturity, and not just at maturity. This is what the terminology *American/European* refers to; it has nothing to do with geography.

□

Example 2.1.6 (a) An investor owns 1 000 shares of HAL, at \$60.00 per share. If the share price drops to \$50.00 over the next three months, this will give a net loss of \$10 000. To *hedge* against this potential decline in prices, the investor buys a put option to sell 1 000 shares in 3 months time at a price of \$55.00 per share. In this way, the potential loss is limited to \$5000 + the initial premium. If, however, the share price rises to \$63.00, the investor will not exercise the option. In that case the investor's net profit will be \$3000 - the initial premium. The investor thus has put a cap on possible losses without restraining the possible gains.

(b) Investor X, still suffering from the ill–effects of a heavy Christmas season, believes strongly that the shares of pharmaceuticals will rise sharply within the next 3 months. She is willing to *speculate*, and back her hunch to the tune of \$10 000. The shares of PharmCor currently trade at \$50.00. If Investor X buys the shares (200 shares) and the share price rises to \$60.00 in 3 months' time, she will have made a profit of \$2 000, whereas if the price declines to \$40.00, her loss will be \$2 000. A 3 month call option to buy 100 PharmCor shares at strike \$53.00 costs \$200. Suppose now that, instead of buying shares, Investor X buys 50 call options instead. If the share price rises to \$60.00, she will exercise the options and buy 5 000 shares at \$53.00 dollars per share. She will immediately sell these at \$60.00 per share. Her profit is therefore

$$5\,000 \times 60 - 5\,000 \times 53 - 50 \times 200 = 25\,000$$

i.e. a profit of \$25 000, instead of just \$2 000. However, if the share price stays below \$53.00, her losses total the full \$10 000.

□

The previous example shows how options can be used as speculative tools that provide *leverage*: Good outcomes become very good, whereas bad outcomes become very bad.

There are many, many different kinds of option contracts. Just on equity, we have the following examples of *exotic options*:

- Barrier options have a payoff that depends on the underlying asset hitting a certain barrier level during the life of the option.

- Asian options have a payoff that depends on the average price of the underlying asset.
- Bermudan options are halfway between European– and American options. they can be exercised at a number of dates between inception and maturity.
- Lookback options have a payoff that depends on the maximum or minimu asset price during the life of the option.
- Rainbow options are options on more than one underlying asset, and have a payoff that depends on the best or worst performing asset in the basket.

In addition, many other securities have option–like qualities:

- *Callable* bonds allow the issuer to call back the bond before maturity. Thus an investor who is long a callable bond is long a standard bond, and short a call option on that bond.
- A *convertible bond* allows the holder to convert the bond to a predetermined number of shares in the issuing company.
- Because of *limited liability*, the shares of a limited company can themselves be regarded as a kind of put option — a put on the value of the company. When the value of the company becomes negative (i.e. when its liabilities exceed its assets), the shareholders are not required to make up the difference. This is why share prices cannot become negative. When the value of a company becomes negative, the shareholders are able to sell their (negative) stake in the company at a strike price of zero.

2.1.3 Structured Products

Financial engineers may package together several financial securities, such as shares, bonds and commodities, together with derivatives to form complex *structured products*. One of the main reasons this is done is to give corporations the ability to borrow at lower cost. Some structured products are designed to appeal to investors: They protect the principal of an investment, while allowing for the possibility of participation in some exciting up–side. Credit risk mitigation is another important source of structured products.

Example 2.1.7 A *collaterized debt obligation* (CDO) is created by pooling together a number of debt obligations (such as bonds or mortgages). For example, many homeowners have funded the purchase of a house by taking a loan from a bank — a mortgage — on which they make monthly payments. There are various reasons why the bank might want to get these mortgages of its books. The credit quality of an individual homeowner is typically not that good — certainly not AAA — and thus it may not be easy for a bank to sell individual loans. However, it is unlikely in the extreme that *all* homeowners will default. In fact, its probably unlikely in the extreme that more than 10% of homeowners will default.

To get the loans of its books, the bank sets an intermediary, a so–called *Special Purpose Vehicle*, which buys the loans from the bank. The assets in the SPV are protected by a firewall: Even if the bank goes bust, its creditors have no claim on the assets of the SPV.

The SPV now pools the mortgages together to form a new security, a collaterized mortgage obligation (CMO). These are sold to investors. The monthly payments made by the homeowners are then sliced into coupon payments going to various tranches. For simplicity,

assume that there are just three tranches, namely *senior, mezzanine* and *junior*. Payments then cascade down the tranches, as follows: First the investors in the senior tranche are paid out. If any money is left over, the mezzanine is paid. Should any money still remain, it goes to the junior tranche (or *equity tranche*[4]). Thus if any homeowners default, the junior tranche takes losses first. Only when the junior tranche is completely wiped out will the mezzanine tranche begin to take losses. And only when the mezzanine tranche is wiped out, will the senior tranche suffer.

Because it is extremely unlikely that more than 10% of the mortgages will default, perhaps 90% of the value of the CMO goes into the senior tranche, with a very high AAA-credit rating. Perhaps another 5% will go into the mezzanine tranche, with maybe a BBB rating, and the remainder into the junior tranche. By this alchemy a collection of individually risky debts is magically tranformed into something that is almost entuirely AAA.

Because of the relation between risk and return, investors in the senior tranche will get paid the lowest rate (e.g. LIBOR +30 basis points). The mezzanine rate is riskier, so gets a slightly better rate (say LIBOR +1%). The junior tranche gets the remainder. In that way it is like *equity*: Shareholders of a company get whatever is left over after the company has paid all its debts.

In the same way, but using bonds, one can construct CBO's (collaterized bond obligations), etc.

Collaterized debt obligations became very popular in the years 2000-2007. Indeed, they were so popular that the supply of actual loans was too small to meet the demand for CDOs. Hence many CDOs were created *synthetically*. A synthetic CDO is constructed from credit derivatives, the so–called *credit default swaps* (CDS), in such a way to have the same (or nearly the same) payoffs as the CDO that it is mimicking.

In the financial press, CDOs and CDSs are widely held to be the cause of the 2007 Credit Crunch and the recession that followed.

□

2.2 Derivatives Markets

Derivatives are traded either *over–the–counter* (OTC) or on organised exchanges. OTC derivatives can be tailored to the specific needs of the customer, whereas exchange–traded derivatives are standardised contracts. One of the difference between forwards and futures is that forwards are OTC–instruments, whereas futures are exchange–traded.

Here's a fairly stupid analogy: If you go to a pharmacy with a prescription from your doctor, you will get the exact dosage of exactly the right medicine for your condition. This is an OTC transaction. Buying cornflakes at the supermarket is different: All boxes are the same, i.e. they are standardized. You cannot buy exactly 253g of cornflakes.

As is stated in all books on derivatives, the growth of the world–wide market in derivatives has been stupendous in recent years, is is not showing any signs of abating. The following tables show the current state of the global market in derivatives:

[4] Also known as the *toxic waste.*

Table 19: Amounts outstanding of over-the-counter (OTC) derivatives

By risk category and instrument

In billions of US dollars

	Notional amounts outstanding					Gross market values				
Risk Category / Instrument	**Dec 2009**	**Jun 2010**	**Dec 2010**	**Jun 2011**	**Dec 2011**	**Dec 2009**	**Jun 2010**	**Dec 2010**	**Jun 2011**	**Dec 2011**
Total contracts	**603,900**	**582,685**	**601,046**	**706,884**	**647,762**	**21,542**	**24,697**	**21,296**	**19,518**	**27,285**
Foreign exchange contracts	**49,181**	**53,153**	**57,796**	**64,698**	**63,349**	**2,070**	**2,544**	**2,482**	**2,336**	**2,555**
Forwards and forex swaps	23,129	25,624	28,433	31,113	30,526	683	930	886	777	919
Currency swaps	16,509	16,360	19,271	22,228	22,791	1,043	1,201	1,235	1,227	1,318
Options	9,543	11,170	10,092	11,358	10,032	344	413	362	332	318
Interest rate contracts	**449,875**	**451,831**	**465,260**	**553,240**	**504,098**	**14,020**	**17,533**	**14,746**	**13,244**	**20,001**
Forward rate agreements	51,779	56,242	51,587	55,747	50,576	80	81	206	59	67
Interest rate swaps	349,288	347,508	364,377	441,201	402,611	12,576	15,951	13,139	11,861	18,046
Options	48,808	48,081	49,295	56,291	50,911	1,364	1,501	1,401	1,324	1,888
Equity-linked contracts	**5,937**	**6,260**	**5,635**	**6,841**	**5,982**	**708**	**706**	**648**	**708**	**679**
Forwards and swaps	1,652	1,754	1,828	2,029	1,738	176	189	167	176	156
Options	4,285	4,506	3,807	4,813	4,244	532	518	480	532	523
Commodity contracts	**2,944**	**2,852**	**2,922**	**3,197**	**3,091**	**545**	**458**	**526**	**471**	**487**
Gold	423	417	397	468	521	48	45	47	50	82
Other commodities	2,521	2,434	2,525	2,729	2,570	497	413	479	421	405
Forwards and swaps	1,675	1,551	1,781	1,846	1,745					
Options	846	883	744	883	824					
Credit default swaps	**32,693**	**30,261**	**29,898**	**32,409**	**28,633**	**1,801**	**1,666**	**1,351**	**1,345**	**1,586**
Single-name instruments	21,917	18,494	18,145	18,105	16,881	1,243	993	884	854	962
Multi-name instruments	10,776	11,767	11,753	14,305	11,752	558	673	466	490	624
of which index products	...	7,500	7,476	12,473	10,466					
Unallocated	**63,270**	**38,329**	**39,536**	**46,498**	**42,609**	**2,398**	**1,789**	**1,543**	**1,414**	**1,977**
Memorandum Item:										
Gross Credit Exposure						3,521	3,581	3,480	2,971	3,912

Figure 2.1: Source: *BIS Quarterly review*, December 2012

Here are some facts to illustrate the growth of derivatives markets: Note that the December 1999 BIS Quarterly Review reported the total amounts outstanding as roughly \$100 trillion. Three years later, it rose to around \$150 trillion, and currently it stands at about\$700 trillion.

Do note however, that the notional amounts (i.e. amounts outstanding) are not the same as the values of the contracts. They are an estimate of an equivalent position in the spot market, i.e. one that has the same response to market factors. For example, a future on a share is approximately equivalent to the share, but the future costs nothing to enter into, i.e. initially has zero market value. Similarly, the premium for a call option is much smaller than the value of the delta–equivalent position in the underlying share.

The amount outstanding for OTC contracts is quoted above as \$648.7 trillion. The gross market value of these contracts amount to “only” \$27.3 trillion (up from \$ 2.8 trillion in 1999). Because exchange–traded instruments are usually marked to market, the value of these instruments at any time is very small.

The corresponding amounts for South Africa are harder to determine. According to data published by the South African Reserve Bank, the notional amounts for OTC derivatives held by banks was R 24 trillion in June 2010. The corresponding figure for exchange traded derivatives was R 800 billion.

Table 23A: Derivative financial instruments traded on organised exchanges

By instrument and location

Notional principal in billions of US dollars

	Amounts outstanding				Turnover					
Instrument / location	**Dec 2010**	**Dec 2011**	**Mar 2012**	**Jun 2012**	**2010**	**2011**	**Q3 2011**	**Q4 2011**	**Q1 2012**	**Q2 2012**
Futures										
All markets	**22,312.0**	**22,924.1**	**24,282.6**	**23,741.6**	**1,380,538.9**	**1,524,140.9**	**412,452.8**	**285,308.7**	**306,691.3**	**312,142.5**
Interest rate	21,013.4	21,718.9	22,900.0	22,414.5	1,235,907.4	1,359,130.8	365,101.7	248,177.1	280,139.8	273,854.9
Currency	170.2	221.2	225.0	214.6	35,771.2	37,627.6	10,527.1	8,103.0	8,120.8	8,687.7
Equity index	1,128.4	984.0	1,157.5	1,112.5	108,860.3	127,382.4	36,824.0	29,028.6	18,430.7	29,599.9
North America	**11,863.5**	**13,107.9**	**14,141.2**	**13,246.9**	**729,195.9**	**822,958.4**	**222,628.4**	**151,434.2**	**174,504.9**	**166,856.9**
Interest rate	11,351.1	12,568.1	13,547.7	12,672.4	658,193.5	740,210.8	198,582.3	132,423.4	160,649.6	147,450.2
Currency	114.8	150.6	149.2	139.7	28,649.0	30,188.7	8,345.6	6,475.3	6,435.1	6,928.4
Equity index	397.6	389.2	444.3	434.8	42,353.4	52,558.8	15,700.5	12,535.6	7,420.3	12,478.3
Europe	**6,345.3**	**6,531.1**	**7,161.3**	**7,426.7**	**533,297.9**	**565,184.7**	**150,755.3**	**105,495.2**	**104,997.1**	**111,332.2**
Interest rate	5,816.6	6,100.0	6,665.3	6,948.4	498,836.1	525,692.3	139,303.5	97,014.4	99,644.5	103,050.7
Currency	2.5	2.7	3.7	2.8	255.2	363.3	99.3	82.8	68.9	106.7
Equity index	526.1	428.4	492.2	475.5	34,206.5	39,129.2	11,352.5	8,398.1	5,283.7	8,174.9
Asia and Pacific	**3,168.6**	**2,339.1**	**2,157.3**	**2,221.8**	**92,273.5**	**107,490.1**	**30,805.3**	**22,091.0**	**20,342.0**	**26,615.1**
Interest rate	2,982.8	2,181.9	1,949.4	2,018.4	60,899.6	71,504.2	20,857.1	13,969.2	14,528.4	17,699.5
Currency	1.5	8.1	8.6	9.0	1,594.4	2,001.9	673.8	418.9	362.3	389.0
Equity index	184.2	149.1	199.4	194.4	29,779.6	33,984.0	9,274.3	7,702.9	5,451.4	8,526.5
Other Markets	**934.7**	**946.1**	**822.8**	**846.3**	**25,771.7**	**28,507.6**	**8,263.9**	**6,288.2**	**6,847.3**	**7,338.2**
Interest rate	862.9	868.9	737.6	775.3	17,978.1	21,723.5	6,358.7	4,770.1	5,317.5	5,654.5
Currency	51.4	59.9	63.6	63.1	5,272.6	5,073.7	1,408.5	1,126.0	1,254.5	1,263.5
Equity index	20.4	17.4	21.6	7.9	2,520.9	1,710.5	496.7	392.1	275.3	420.2
Options										
All markets	**45,634.6**	**33,639.2**	**40,697.6**	**35,789.7**	**606,661.8**	**635,363.3**	**184,400.1**	**132,621.7**	**120,373.7**	**116,365.1**
Interest rate	40,930.0	31,579.6	37,946.9	33,231.2	468,872.0	466,281.3	139,524.5	97,614.9	98,797.1	83,221.6
Currency	144.2	87.2	110.1	110.9	3,048.5	2,525.1	688.5	519.8	573.9	618.8
Equity index	4,560.3	1,972.4	2,640.6	2,447.6	134,741.3	166,556.9	44,187.2	34,487.0	21,002.7	32,524.7
North America	**24,353.4**	**18,025.8**	**20,178.1**	**16,712.9**	**261,543.9**	**262,514.8**	**78,400.1**	**53,926.3**	**62,060.8**	**52,262.5**
Interest rate	22,070.2	17,779.1	19,788.3	16,252.3	225,342.9	218,368.1	65,025.4	43,082.0	55,449.4	40,989.3
Currency	72.3	49.1	70.1	76.6	1,600.9	1,510.4	401.6	323.9	354.2	418.5
Equity index	2,210.9	197.6	319.7	384.0	34,600.2	42,636.4	12,973.0	10,520.4	6,257.2	10,854.7
Europe	**19,247.2**	**14,280.8**	**19,173.8**	**17,505.5**	**251,485.3**	**258,265.4**	**78,432.2**	**57,197.9**	**43,610.5**	**42,617.4**
Interest rate	17,320.8	12,879.8	17,446.3	15,970.1	233,930.0	237,541.6	72,325.5	52,636.5	40,935.0	38,971.2
Currency	0.3	0.3	0.2	0.2	5.1	3.4	0.8	0.8	0.6	1.1
Equity index	1,926.1	1,400.6	1,727.4	1,535.2	17,550.2	20,720.4	6,105.9	4,560.6	2,674.9	3,645.1
Asia and Pacific	**383.3**	**349.6**	**544.4**	**486.9**	**82,757.4**	**102,912.5**	**24,940.0**	**19,466.1**	**12,569.7**	**18,754.9**
Interest rate	3.5	15.6	6.5	3.0	2,605.3	2,719.3	659.7	791.9	859.0	976.5
Currency	0.3	0.8	1.2	1.3	6.2	257.0	103.6	51.1	50.8	50.6
Equity index	379.4	333.3	536.7	482.6	80,145.8	99,936.1	24,176.7	18,623.1	11,659.9	17,727.8
Other Markets	**1,650.7**	**983.1**	**801.3**	**1,084.4**	**10,875.2**	**11,670.6**	**2,627.8**	**2,031.4**	**2,132.7**	**2,730.3**
Interest rate	1,535.5	905.1	705.9	1,005.8	6,993.8	7,652.3	1,513.8	1,104.5	1,553.7	2,284.5
Currency	71.3	37.0	38.7	32.7	1,436.3	754.3	182.5	143.9	168.4	148.7
Equity index	43.9	40.9	56.8	45.9	2,445.1	3,264.0	931.5	782.9	410.6	297.1

Figure 2.2: Source: *BIS Quarterly review*, December 2012

2.3 Tales of Woe

Derivatives provide high leverage at low cost: A cheap initial position can lead to tremendous profits and losses. It is not surprising, therefore, that the misuse of derivatives as hedging or speculative tools can lead to disaster. In this section we will indulge in a little gossip and *schadenfreude*, in the guise of cautionary tales.

Most of the stories will concern financial and industrial corporations. However, the individual small investor is very much at risk: Take the story of the small investor who had written *naked puts*[5] on the FTSE 100. In the Black Monday crash of October 1987, the buyers of put options made tremendous profits, i.e. the writers of puts made huge losses. Though the small investor's brokers tried to get hold of him to close out his position, he could not be reached. He could not be reached because he was 13 years old, and at school. By the time he had been located, he owed more than £1 million. Of course, he could not pay.

We present here just a small selection of the famous crash-and-burn stories which form part of the derivatives culture. There are many, many more, and these days any financial mishap is initially blamed on the use of derivatives, e.g. the recent case of Enron (where derivatives played hardly any role).

2.3.1 Orange County

Orange County is wealthy district in California. The events related here occurred at a time when a 69–year old man called Robert Citron was the County Treasurer, and he had been on the job for over twenty years. Citron's returns were well above those of most other municipal funds, which is why he was re–elected time and again. He accomplished these high returns by taking highly risky positions on interest rates, and he was usually lucky. As usual, no one cares to scrutinise too closely the position of a successful trader, and admittedly Citron was under pressure to perform: the electorate kept voting lower taxes for themselves, so the deficit had to be made up from investment income. Although there were ample signs that Citron's fund was dangerously leveraged, the electorate did not care. Thus, in 1994, this same electorate was left with a bill which totalled \$1 000 for every man, woman and child in the county. This was the biggest money market disaster on record at the time, leading to losses of about \$1.7 billion and the axing of 2 000 county jobs.

Citron's mail–order psychic had predicted that interest rates would go down, and Citron believed this with the fervour of a religious fanatic. Asked why he was convinced that rates would rise, he replied: "I am one of the largest investors in America. I know these things." To increase the leverage in his bet, Citron therefore invested \$8 billion in *inverse floaters* — bonds whose coupons go up if rates go down. Because (1) the coupon increases in size, and (2) coupon payments are discounted at lower rates when rates go down, inverse floaters are far more sensitive to interest rate movements than ordinary bonds and floating rate notes. He also invested in range notes, which only pay a coupon if interest rates stay within certain bounds. By using reverse repos, he managed to borrow heavily to take further bets on falling rates.

Until 1994, Citron did very well indeed: He made \$750 million between 1991 and 1994, and was one of the darlings of Wall Street, dealing with Merrill Lynch, Credit Suisse, J.P. Morgan and Goldman Sachs. Merrill Lynch, for example, made over \$100 million from Citron's

[5] Recall that the writer of a put option is obliged to buy shares at a certain price. If the writer isn't already hedged, e.g. by being short the share, or by being long call options, the put option is said to be *naked*.

deals in fees and commissions over 1993–1994. When the Federal Reserve increased rates in early 1994, to everyone's surprise, the fan was hit by a substance with a less–than–pleasant odour, and Citron's portfolio began to collapse. Merrill Lynch, who had been intermittently worried about excessive risk since 1992 (...oh, but those fees and commissions...) now became seriously worried, and tried in earnest to devise a strategy for Orange County. Citron, however, refused to play ball, claiming that Merrill Lynch had not informed him about the risk inherent in the instruments he had bought. Orange County declared bankruptcy, even while they were still in the black, in order to protect its assets, but this attempt was largely unsuccessful. Merrill Lynch was exonerated, but Citron was arrested. In order to ward off a possible 14 years in jail, the big omniscient investor declared: "In retrospect I was not the sophisticated treasurer I said I was." Psychologists (for the defense?) claimed that Citron had the mathematical ability of a 13–year old, and that he belonged to the bottom 5th percentile in his ability to think. Because he was stupid, he only got one year of community service, although the mail–order psychic is probably still on the loose.

At the time, S&P rated Orange County as AA, while Moody's assigned its debt Aa1[6]. These high credit ratings imply a very small probability of bankruptcy and default.

2.3.2 Metallgesellschaft

Metallgesellchaft is large German conglomerate, one of the country's foremost commodities groups. In 1993 the company was an ailing metal refining group. Pressurised by Deutsche Bank, one of the company's major share holders, MG had taken on DM 9 billion in debt in the hope turning the company's fortunes, but nothing seemed to work. The one ray of hope was provide by its US subsidiary, MG Refining and Marketing. MGRM aggressively sold long–term forward contracts on oil to its clients, promising them fixed prices for up to 10 years. As a result, MG was exposed to rising oil prices. One of their traders, Arthur Benson, came up with the idea to hedge their exposure by buying 3–month oil futures on the NYMEX, rolling them over as they neared maturity. If oil prices rose, the loss on the forward contracts would be balanced by the gain on the futures.

Benson has been described as a market animal. In the mid–1980's he was the defendant in a market–related lawsuit, and his own lawyers described him as "caught up in an addiction to the market.... [He] became, with regard to the market, irrational." Benson had worked for MG before, but was fired after losing $50 million dollars. When the Gulf War sky–rocketed his massive long jet fuel futures portfolio by $500 million, MG wanted him back. Irrational market junkies are always in demand as traders.

Towards the end of 1993, the oil price began to fall. One would think that this was no problem for MG, since they were worried about rising prices. However, one of the important differences between forwards and futures is that forwards are OTC, whereas futures are exchange–traded and marked to market. This means that cash flows on the forward contracts would only have materialized when the forwards matured, up to 10 years in the future. On the other hand, the marking–to–market process for futures means that losses on the futures contracts would have to be settled daily. As we have already stated, if oil prices rose, then the loss on MG's forward contracts would be balanced by the gain on the futures. But it is also true that if oil prices decreased, the gain on the forwards would be balanced by the loss on the futures. A loss on a futures position would have to be settled immediately however,

[6] Credit ratings are discussed in Section 5.5

whereas the corresponding profit on the forwards would only be received as cash much later.

MG ran into liquidity problems. Benson, after he was fired for the second time, claimed that his bosses had ordered him to illegally corner the US heating oil market, so MG's position must have been huge (whether Benson was lying or not). When oil prices fell, MG was responsible for large margin calls to settle the losses on the futures. They simply ran out of cash, and management's response was to close out all positions, leading to losses of \$1.3 billion.

Deutsche Bank claimed that the losses were due to a single *rogue trader*: Arthur Benson. Others, however, claim that MG and Deutsche Bank failed to understand the nature of the hedging strategy; after all, the losses on the futures would eventually be recouped by the forward position. When management cancelled the forwards, they were giving away \$1.3 billion.

The truth is probably somewhere in between. As we shall see, forwards and futures differ in other respects. In the story of Metallgesellschaft, the futures market was initially in *backwardation* (where short–term forward prices are higher than long–term forward prices), and then moved to *contango* (where short–term forward prices are lower than long–term forward prices, the usual state of affairs). While the market was in backwardation, rolling over the futures contracts led to gains, but when it moved to contango, Benson was losing \$20 to \$30 million a month. There is less–than–perfect correlation between forwards and futures prices, which makes real losses possible. This risk, *basis risk*, is difficult to model or to hedge. It is doubtful, however, that losses would have been anywhere near as severe if management had not intervened, and the whole episode proved very embarrassing to Deutsche Bank, who was at that time trying to establish itself as a major derivatives player.

2.3.3 Barings

> *A glance at the proprietary trading department of a US bank would have shown Baring's top brass how inadequate Leeson was for the job — indeed, how inadequate many of them were too. The top US bank dealing rooms are stuffed with university professor, dealers with training in higher mathematics, managers with long experience of derivatives. But they didn't look. The only thing they seemed to see is that the Americans made big money from proprietary trading and paid themselves unimaginably large bonuses. They wanted that too but did not bother to find out that the whole process required a degree of professionalism they did not posses.*
>
> —- Richard Thompson
>
> *Apocalypse Roulette*

Barings Bank was Britain's oldest bank, founded on Christmas Day 1762 – bankers to the Queen. It was a small, conservative and risk averse establishment, but had to join the real world when British markets were deragulated in 1986 (the so called *Big Bang*). In order to compete with the influx of American banks and brokers, Barings bought a small broking firm and started trading Japanese warrants. The Japanese operation made a fortune for Barings, and established them as a major player in the Asian markets.

In 1992, Barings appointed Nick Leeson to its Singapore office. They needed a back–office clerk to process their arbitrage operation between the Osaka (Japan) and Singapore markets, as well as a floor manager for their dealing with the Singapore International Monetary Exchange (SIMEX). Leeson got both jobs.

This *wunderkind* has been described as a "turbo–arbitrageur." Within a year, he had earned £10 million, about 10% of the total profits of Barings Bank. Over the next few years, he would, on paper, be responsible for 20% of total profits. Such high returns should have alerted management to the risky positions he was entering. As general manager of the Barings Futures subsidiary, however, Leeson controlled both the trading and back office functions. This allowed him to effectively conceal the riskiness of his trades. The separation of trading and back–office functions is one of the cardinal rules of effective risk management, but Barings was, after all, a gentleman merchant bank, and its top management understood very little about the dangers of derivatives trading.

In fact, though Leeson was making large profits on paper, he was actually sustaining heavy losses as well, but managed to hide these in the infamous error account 88888. Before he was even registered as a trader on Simex, he began trading Nikkei 225 futures, promptly lost £40 000, and hid it in 88888. By the end of 1992, 88888 contained £2 million in undisclosed losses. By the end of 1993, he had lost £25 million.

By January 1994, Leeson had shorted £30 million's worth of options on the Nikkei 225, to cover the losses. He shorted both calls and puts, and the resulting position, called a *short straddle* would have been profitable if prices remained stable. Nevertheless, this was outright speculation, but Barings believed he was arbitraging the Nikkei 225 between Singapore and Osaka. In February 1994, this option portfolio was valued at ¥2.8 billion. Barings sent an internal auditor to investigate these massive profits, who reported on the weakness of internal controls and recommended the separation of trading and back office activities. He also recommended that every effort should be made to ensure that the star trader wasn't poached by the competition. Barings responded by giving Leeson a £450 000 bonus, and assigning part-time responsibility for back office activities to financial manager from Hong Kong.

In January 1995, the Nikkei 225 dropped by about 1000 points after an earthquake measuring 7.2 on the Richter scale demolishes the city of Kobe. The straddles shorted by Leeson are now deep in the money. Singlehandedly he tried to force the Nikkei back up, buying futures like a maniac. This double–or–nothing strategy failed, leaving him even deeper in debt. On the 23$^{\text{rd}}$ of February, Leeson lost £144 million, and decided to call it quits. Fleeing first to Kuala Lampur, he was finally picked up by police in Germany.

Barings lost $1.3 billion. As its assets amounted to just about $0.9 billion, it went bankrupt in March 1995, and was bought by the Dutch group ING for a princely £1.

And what about Nick? He was sentenced to $6\frac{1}{2}$ years by a Singapore court, having already languished some months in a German jail. He was released early, in July 1999. Three months later, Leeson was paid $100 000 for an appearance, but Leeson is still in hock to Barings' creditors (to the tune of $160 million), and gets to keep only 35% of the money he makes on speeches. These days he fashions himself as a role model:

AIB TRADER MISSING

Ex-Barings trader Leeson says 'no lessons learned'.

Posted: 11:23 AM (Manila Time) — Feb. 07, 2002

By Agence France-Presse

LONDON - Nick Leeson, the rogue trader who brought down Barings bank in 1995, said Thursday that "nothing has been learned" after a dealer with Ireland's biggest bank AIB vanished having allegedly run up losses of 750 million dollars. "I am shocked nothing has been learned from my case and the same thing has been allowed to happen," said Leeson, writing in the Mirror, a British tabloid.

"It's staggering that financial security at these huge firms is so lax. Where are the checks and regulatory audits? There has been chapter and verse written about my experience at Barings but all that counts for nothing."

How noble of him...

According to that font of all knowledge, Wikipedia, Leeson these days earns his keep by advising corporations on risk management (presumably "What not to do") and by writing books (*Back From The Brink: Coping with Stress*, Virgin Books, 2010). That sounds a lot like a self–help book for rogue traders. Jérôme Kerviel (a trader at Sociéte Générale who lost \$6.5 billion in 2008), and Kweku Adoboli (who lost \$2.3 billion while trading for UBS in 2011) have no doubt eagerly devoured this important contribution to the genre as they languish in jail.

2.3.4 Long-Term Capital Management

Markets can remain irrational longer than you can remain solvent.
— John Maynard Keynes

...the market will find you...

— Marek Musiela

LTCM was a *hedge fund* founded in 1994 by John Meriwether, who had been one of Salomon Brothers' star bond traders. He recruited a stellar cast of traders and academics, including Myron Scholes and Robert Merton, who were to win the Nobel Memorial Prize in Economic Science for their work on the Black–Scholes model in 1997. David Mullins, a former vice–president of the Federal Reserve Board, was also on the team, as were a string of recent hotshot Ph.D's. Dazzled by this dream team, who promised annual returns of about 30%, investors scrabbled to hop on the band wagon, coughing up hundreds of millions of dollars. By the time LCTM was ready to run, it had corralled \$1.25 billion in funds to invest.

The "hedge" in "hedge fund" is quite misleading. It refers to the fact that hedge funds are allowed to take positions which are market neutral, i.e. both short and long positions are

allowed. However, this is not at all the same as hedging. LTCM specialized in *convergence trades*. For example, dollar denominated emerging market bonds have higher yields than government bonds. If you short 5–year government bonds and use the resulting cash flow to buy 5–year emerging market bonds, the net cash flow today is zero. The income you receive from the emerging market bonds is higher than the payments you owe on the government bonds, so if you take a large enough position you will have made a neat pile of money after 5 years — and this at no initial net cost.

If this seems a clever way to make money for nothing, think again. There is a reason that emerging market debt has a higher yield than sovereign debt: Your emerging market bonds are far more likely to default. Should that happen, your emerging market bonds will be practically worthless, but you will still be responsible for the payments on the government bonds. This market neutral position therefore contains a substantial amount of credit risk. Even on the short term, the position is very risky, as an increase in the yield spread will lead to large short–term losses. Nevertheless, the principle is sound, and generally far less risky than taking an outright position. LTCM used sophisticated models to ensure that its long– and short positions were highly correlated. They did not place bets on any outright event (such as increasing interest rates). Instead, they (1) bet on convergence between US, Japanese and European sovereign bonds; (2) bet on convergence between of–the–run and on–the–run sovereign bonds; (3) bet on convergence between German and other European sovereign bonds; and (4) took long positions in emerging market debt, backed by shorting Treasuries. They attempted to have a risk profile equivalent to investing in the S&P 500 (and boasted that they were having trouble getting the risk up high enough).

Since yield spreads are tiny, huge bets had to be placed in order to generate profits. This could be done. Hedge funds were hardly regulated by the SEC, and were often highly secretive about their positions. By using short positions to fund long positions, collateralized by repos, LTCM were able to leverage their fund by a large factor. Because of their reputation and connections, banks fought to do business with them, and they got the sweetest deals. Right from the start, LCTM refused to pay haircuts or insisted on substantial reductions. Normally if you borrow a security, you have to post a little more collateral than the value of the security, e.g. \$1.01 million collateral for a \$1 million Treasury bond, and more on a credit risky bond. This difference is known as a haircut. Haircuts therefore limit how much you can trade. However, at one stage LTCM paid only \$500 million in haircuts for debt amounting to over \$100 billion.

Initially, all went well. In 1994, LTCM earned 28%, although their fees were so high that investors only saw 20%. LTCM charged 2% p.a. plus 25% of profits. Other hedge funds typically charged 1% and 20% respectively, while mutual funds charge 1.4% and 0%. In 1995, LTCM earned a phenomenal 59% (43% after fees), and in 1996 they followed with a 57% (41%) return. At this time, LTCM began to find it difficult to locate lucrative convergence–arbitrage deals — in fact, they claimed that convergence had happened much faster than they had expected, so that returns had far exceeded management's expectations. Accordingly, they predicted that returns in 1997 were likely to be much lower; indeed, they managed only 25% (17%), which was weak compared to the S&P 500's return of 31%. Still, investors got back \$1.82 for every \$1.00 invested over the 4–year period, and LTCM's partners earned \$1.5 billion in fees.

By now, word on bond convergence–arbitrage was out, and the competition was intense. Citing lack of opportunity, LTCM soured relations with some investors by returning their funds, so that they could increase the leverage of their fund. Very soon, they were to wish

that they hadn't, because the markets went out of control.

At the beginning of 1998 they controlled \$4.7 billion in equity and had borrowed \$125 billion. It had a huge position in interest rate swaps, with notional amounts outstanding amounting to \$1.25 trillion — a staggering 5% of the world market.

LTCM's positions were designed to perform under conditions of relative stability — conditions which are necessary for convergence. For example, they shorted options on the S&P 500, thus selling volatility. This would lead to profit if volatility remained were it was. When markets go mad, however,volatility goes up. Options have a greater chance of being exercised, causing losses for the short side. LTCM sold options at 19% volatility. As these were exchange–traded, LTCM was responsible for margin calls on losses.

Investors ditch credit risky and illiquid bonds, and buy quality instead. In this *flight to quality*, yield spreads don't converge; they increase. The trouble started in May 1998, when MBS spreads went up. Bond arbitrageurs began a cycle of selling to reduce their exposure, including emerging market debt. Not good for LTCM, and it also compounded a crisis in the Asian markets. Indonesia, in particular, was a trouble spot, as currency speculators engineered a currency devaluation. The IMF bailout failed, riots ensued, and Suharto was forced to resign after 32 years in power.

Worst of all, LTCM had bet that spreads between Russian and Japanese bonds would decrease, i.e. that Russian yields would fall and Japanese yields would rise. After Indonesia, however, Russia was targeted To prevent a devaluation of the rouble, the Russian central bank had tripled interest rates by the end of May. Implied volatility on the S&P 500 rose to 27%. By the end of June, LTCM had lost 14% of its value.

On 17 August 1998, Russia devalued the rouble and defaulted on its rouble–denominated sovereign debt. On the 21st of August, the Dow fell by 280 points before noon, and then recouped its losses, a tremendous surge in volatility. Credit spreads sky–rocketed, and it has been reported that LTCM lost \$500 million on that day alone. By the first of September, their equity base was down to \$2.3 billion, having lost 52% for the year–to–date, and their leverage had increased from 27–to–1 to 55–to–1. They were rapidly running out of cash, and had sold their most liquid assets. On the 2nd of September offered investors a special opportunity to invest at reduced fees, but there were no takers. By now, spreads had widened to the largest since the October 1987 crash. LTCM partners believed that spreads would eventually have to narrow, and they saw this as a golden opportunity, if only they could find the cash to invest. But they had great difficulties just meeting payments on their losses. Everything seemed to move against LTCM: Why would Britain experience a flight to quality, while German investors remained calm? It did not make sense. It is likely that, despite the secrecy of their positions, were singled out for attack.

By the 22nd of September, equity was down to \$600 million, pushing the leverage even higher. The next day, the Federal Reserve organized a bailout, injecting \$3.6 billion and taking over control of the company. It's not the Fed's job to rescue ailing banks, but it was feared that LTCM's position could cause seizures in world markets. The Fed's chairman, Alan Greenspan, had the following to say: *"Had the failure of LTCM triggered the seizing up of markets, substantial damage could have been inflicted on many market participants, including some not directly involved with the firm, and could have potentially impaired the economies of many nations, including our own.*

LTCM's partners lost everything. Barely a year later, however, Meriwhether had managed to raise \$250 million for a new venture. The market had found him, stayed irrational just long enough to bankrupt him, and then promptly forgot all about it. After all, *this* time it

would be different. . . .

Chapter 3

Forwards and Futures

3.1 Futures Contracts

3.1.1 Futures Contracts vs Forward Contracts

The most common financial transactions are (still) spot transactions, which are agreements to buy/sell *today* for immediate physical delivery. For example, grain farmers historically sold grain at a centralized location (a market) to sell their grain. Manufacturers and producers will often be able to tell in advance that they will need to buy or sell require a certain quantity of goods at some future date, however. For example, a grain farmer knows that she will need to sell her grain at the end of the season, and a miller knows that he will need to buy grain to mill, which he must then sell to a baker. Forward contracts are financial instruments which allow counterparties to enter into *forward transactions*, agreements to buy or sell at a future date. We recall the definition of a forward contract:

Definition 3.1.1 A *forward contract* is an OTC (private) agreement between a buyer and a seller to exchange a specified quantity of asset at a specified price for delivery at a specified place and time.

□

The party who agrees to buy is said to have a *long* position, and whereas the party who agrees to sell has a *short* position. The prespecified price of the asset is called the *delivery price* and is set such that the initial value of the forward contract is zero, so that no cash is exchanged when entering the contract. Thus a forward contract *initially has zero value*. As time progresses, however, the contract will come to have a positive or negative value. The prespecified date is known variously as the *delivery date*, the *expiry date* or the *maturity* of the forward contract.

A *futures contract* is very much like a forward contract: It is an agreement to accept or deliver a specified asset at a specified time for a specified price. The mathematics of forward– and futures contracts is therefore very similar. For example, the initial value of a futures contract is zero, just like for a forward contract. Nevertheless, there are important differences:

- Forward contracts are traded OTC (over the counter). This means that they are issued by a financial institution for a specific client, i.e. they are be tailored to individual

parties' needs. Futures, on the other hand, are traded on an exchange. To facilitate trading, they have *standardized features*. The exchange specifies what can be delivered, where it can be delivered and when it can be delivered. Thus contract sizes are fixed, and there are only a few grades (quality) and expiry dates available.

- While it is obvious what the advantages of having a contract tailored to individual needs are, the advantages of using futures may need spelling out. Futures are exchange traded, and therefore far more **liquid** (easy to convert to cash). Generally an investor who is long porkbelly[1] futures has no intention of ever accepting delivery. In 95 – 99% of cases, the contract is *closed out* before delivery date. For example, all an investor long a March CBOT 100—ounce gold future needs to do to close out his position is to enter an offsetting short March CBOT 100–ounce gold futures contract. *The vast majority of futures contracts are closed out before expiry, and do not lead to delivery.* An OTC tailor–made contract may be much harder to get out of. Thus forward contracts are generally settled by delivery.

- With forwards, both parties are exposed to *counterparty risk*, the risk that one of the parties is unable to meet its obligations. Futures contracts, however, are settled through a *clearing house* which acts as a middleman. The clearing house is essentially a counterparty for each contract, and will guarantee that a contract will be honoured. This minimizes credit risk for an investor in futures, because the clearing house generally has a high credit rating. To protect itself from defaulters, the clearing house has put mechanisms into place to enable it to detect likely defaulters early on. This is achieved by *marking–to–market* on a daily basis. All parties have a *margin account* at the exchange, and are required to deposit money into it on a daily basis if prices move against them, and from which they can withdraw if price moves are in their favour.

- While it is nearly impossible to regulate forward contracts, futures markets can be regulated.

To summarize:

Forward Contracts	**Futures Contracts**
Traded OTC	Exchange–traded
Tailored (Negotiated)	Standardized grades and quantities
Delivery date negotiated	A small range of delivery months
Settled at end of contract	Settled daily by marking–to–market
Usually settled by delivery	Usually closed out before delivery
Two parties with opposite interests	Between party and clearing house
Unregulated	Regulated

Forwards vs. Futures

Some exchanges also allow one to trade options. On the JSE SAFEX (South African Futures Exchange), there is a hybrid instrument, namely the *Can–Do* option, that "attempts to give investors the advantages of listed derivatives, with the flexibility of OTC contracts." Can–Do's are negotiated by two counterparties, but then valued and marked–to–market by the exchange.

[1]Porkbelly futures used to be an icon on the Chicago Mercantile Exchange (CME), featuring e.g. in the movie *Trading Places*. They ceased to trade in 2011.

Month	Settl	Chg	Open	High	Low
Feb 2013	85.975	0.275	85.6	86	85.375
Apr 2013	88	-0.275	87.8	88.1	87.1
May 2013	95.2	0	95.2	95.2	95.2
Jun 2013	96.9	-0.375	96.65	96.9	96.05
Jul 2013	96.6	-0.4	96.6	96.65	95.85
Aug 2013	96.025	-0.45	95.7	96.05	95.4
Oct 2013	85.975	-0.4	85.85	86.05	85.4
Dec 2013	83	-0.225	82.75	83	82.3
Feb 2014	85	-0.1	85	85	84
Apr 2014	86.3	0	86.3	86.3	86.3
May 2014	91.9	0	91.9	91.9	91.9
Jun 2014	93	0	93	93	93

Figure 3.1: Lean Hog Futures Quotes on the CME, 23 January 2013

3.1.2 Futures Prices and Quotes

Recall that the *forward price* of an asset is the price agreed upon in a forward contract, and not the price of entering a forward contract. Similarly the *futures price* of a futures contract is the agreed upon price for delivery of the underlying asset (grade and quantity) for a particular delivery month. It is *not* the cost of entering a futures contract. In fact, one should perhaps speak of a futures quote rather than a futures price. The futures price changes daily, and it is *changes* in the futures price that must be paid — not the futures price itself. We will shortly discuss this in more detail. In the mean time, here is an example:

The above are quotes for lean hog futures traded on the Chicago Mercantile Exchange (CME) on the 23rd of January 2013. Each contract is for the delivery of 40 000 pounds.

- The SETTLE column contains the settlement price, determined by the exchange. It is an average of the prices at which the contract is traded just before the end of the end of the day's trading. The settlement price is used to calculate daily gains and losses, and determines the amount in an investor's margin account.

- The CHANGE column contains the change from the previous day's settlement price. Thus, for example, the settlement price of the February contract changed by +0.275 cents per pound. An investor who is long one February futures contract will find that the balance in his margin account has increased by \$110.00 (= 40 000 × 0.275 c). An investor who is short a February contract, will find his balance decreased by the same amount. If the balance drops below a certain level, the investors broker may make a *margin call*. We will discuss margins in more detail below.

- The OPEN column refers to the futures price of the first traded contract of the day. For the February 2013 contract, the open price was 85.6 cents/lb. Generally this will be different from the previous day's settlement price.

- The HIGH and LOW columns refer to the maximum and minimum futures prices for that day.

- Sometimes shown, but not shown here, is an INT. column, which refers to the *open interest*, i.e. the total number of contracts outstanding. According to CME data there were 236 705 lean hog futures outstanding on 23 January 2013.
- Also not shown is an *VOL* column, which refers to the estimated volume of trading, i.e. the total number of contracts traded that day. For February 2013 futures, the estimated number of contracts traded was 61 290.

3.1.3 Mechanics of Futures Trading

We briefly discuss *standardization* of futures contracts, the use of *margin accounts* and give an example of the events occurring during the life of a typical futures contract.

Standardization

Each contract will specify some or all of the following:

1. Contract size, e.g. 5 000 bushels, 100 troy ounces, 40 000 lbs., 1 000 barrels
2. Units of quotation, e.g. dollars per barrel for crude oil futures, cents per pound for pork belly futures.
3. Tick size: this is the minimum that a quoted futures price can change by, e.g. $0.01 per barrel for oil futures, 0.00025 cents per pound for lean hog futures.
4. Contract month, e.g. February 2003.
5. Trading hours: CME Lean hog futures are traded 9.05am – 1.00pm central time on the open outcry trading floor on the CME on regular trading days. On the last trading day, however, trading times are 9.05am – 12.00pm.
6. Last trading day: the last day that the contract may be traded. For lean hog futures this is the tenth business day of the contract month.
7. Deliverable grades. In the case of lean hog futures, no hogs will be delivered, but all contracts are cash settled.
8. Mechanics of delivery: This includes the first and last day that notice of delivery can be given, and where delivery can be made.
9. Daily price movement limits: In most cases, the futures exchange also imposes a maximum daily price fluctuation. In the case of gold, for example, the maximum daily price move is $50 per ounce above or below the previous day's settlement price. For lean hog futures, it is $0.03 per pound above or below the previous day's settlement price (equivalent to $1 200 per contract). Trading normally ceases for the day if a *limit move* is made. These price movement limits are set to prevent "hysteria" and "speculative excesses." Whether these attempts to control market volatility are a good thing is controversial. The exchange may in any case opt to increase movement limits, or may abolish them altogether if, for example, limit moves are made on a number of successive days. Lean hog futures have no daily price movement limits on the last two days of trading.

10. Position limits: The maximum number of contracts that a *speculator* may hold. Hedgers are exempt from this. Position limits are imposed to prevent any one individual from cornering the market. For lean hog futures, it is 4 150 futures in any contract month. For spot contracts — within the last two trading days — the limit is 950 contracts.

Settlement and Margin Accounts

Not everyone is allowed to trade futures directly on a futures exchange. Instead, the exchange allows only certain members of the *clearing house* to trade. Investors who wish to trade futures must place orders with their brokers, and these brokers must, in turn, place the orders with a member of the clearing house.

Suppose an investor wants to buy or sell a futures contract. Her broker will insist that she deposit an amount into a margin account. This amount is called the *initial margin.* Its purpose is to act as insurance that the investor is capable of paying for any losses she may incur. Moreover, the amount in the margin account may not drop below a certain level. If it does, a *margin call* is made, and the investor is required to top up her margin account to the level of the initial margin. Suppose, for example, that on 8 January, the investor enters a March futures contract to buy 100 troy ounces of gold on a day that the futures price is \$380.00 per ounce. The initial margin might be set at \$2 000, and the maintenance margin at \$1 500. If at the end of the day the settlement price has dropped to \$376.00, the investor has made a loss of \$400, and the amount in her margin account is now \$1 600. Since the investor pays for her losses immediately, the price for contract delivery must change as well. Thus, at the end of the day, she has paid \$400, and now has a futures contract with delivery price \$376.00. The value of her futures contract is now again zero.

If a day later the settlement price is \$374.00, the investor has lost another \$200, and her margin account now contains only \$1 400. As this is less than the maintenance margin, her broker will make a margin call for \$600, so that the margin account is topped up back to \$2 000. Should the investor be unable or unwilling to comply with the margin call, her broker will close out her position.

If prices move in the investor's favour, she is entitled to withdraw any amount above the initial margin from the margin account. For example, if on the first day the settlement price had increased to \$383.00, she would have been entitled to withdraw \$300.

This daily settlement of profits and losses is called *marking to market.* At the end of each day, the losses (gains) are subtracted (added) to the margin account, thus bringing the value of the contract back to zero.

Minimum margin levels (initial and maintenance) are set by the exchange, but individual brokers may well require greater amounts to be posted. The level of margin depends on the volatility of the underlying asset: Assets whose prices are highly volatile would require more margin than low volatility assets. Maintenance margin is usually about 75% of initial margin. Margin levels are also affected by the objectives of the trader. For example, a speculator may be required to post more initial margin than a hedger.

Interest is usually paid on margin accounts. Moreover, instead of cash, equity or bonds may be posted as margin, though not at full value.

The mechanics of futures trading are best illustrated by example:

Example 3.1.2 On the morning of the 28th of January, investor X buys two July lumber futures contract on the Chicago Mercantile Exchange. The futures price is 261.20. The

contract specifications are as follows:

Trading unit	110 00 board feet of nominal 2×4s in random length from 8 ft to 20 ft.
Quotations	dollars per thousand board feet
Tick size	\$0.10 per thousand board feet = \$11.00 per contract.
Contract months	Jan, Mar, May, Jul, Sep, Nov.
Trading hours (central time)	9.00am – 1.05pm Last day: 9.00am – 12.05pm
Last trading day	Business day immediately preceding 16th calender day of contract month.
Limits	\$10.00 per thousand board feet above or below previous day's settlement price.

Minimum margin requirements, as specified by the CME, are as follows:

	Initial	Maintenance
Speculator	\$1 500	\$1 000
Hedger/Member	\$1 000	\$1 000

Investor X is a speculator. He is therefore required to post initial margin of $2\times$ \$1 500 = \$3 000 with his broker (assuming the broker's margin levels are the minimum levels permitted by the exchange).

The following table shows the movements of daily settlement prices, and the resulting cash flows. Note that each \$0.10 change in the futures price is equivalent to an \$11.00 change in the value of the futures contract, and thus a change of \$22.00 in the value of two futures contracts.

Day	**Settlement price**	**Daily gain**	**Cum. gain**	**Margin account balance**	**Margin call**
28 Jan	262.80	+\$352	+\$352	\$3 352	
29 Jan	260.50	-\$506	-\$154	\$2 846	
30 Jan	258.30	-\$484	-\$638	\$2 362	
31 Jan	254.10	-\$924	-\$1 562	\$1 438	\$1 562
3 Feb	255.60	+\$330	-\$1 232	\$3 330	
4 Feb	256.30	+\$154	-\$1 078	\$3 484	

On the first day, prices move in investor X's favour, and by the end of the day the futures price is up by \$1.60. This implies a gain of $16 \times 11 =$ \$176.00 per contract, and thus a gain of $2 \times 176 =$ \$352.00 for investor X. From here on it is downhill: On the 31st of January, the balance in the margin account drops below the maintenance margin level of \$2 000. Investor X is required to top up the margin account to the initial margin level of \$3 000. The next day is better, and investor X gains \$330. The day after, the margin account contains \$3;484. Investor X can withdraw \$484, if he so wishes. Instead, investor X decides to close out his position by selling two July lumber futures contracts, resulting in a cumulative loss of \$1 078.

□

The next example shows how futures contracts can be used to hedge.

Example 3.1.3 On 16 March, a company knows that it must purchase 100 000 barrels of oil some time in October. The company is exposed to the risk that oil prices go up. It therefore enters 100 November futures contracts on NYMEX, where each contract is for 1 000 barrels of sweet crude oil. On 16 March, we have the following:

$$\begin{aligned} \text{Spot Price} &= \$85.50/\text{barrel} \\ \text{November Futures Price} &= \$84.20/\text{barrel} \end{aligned}$$

On 20 October, the company closes out the contracts. We then have

$$\begin{aligned} \text{Spot Price} &= \$90.10/\text{barrel} \\ \text{November Futures Price} &= \$90.07/\text{barrel} \end{aligned}$$

Ignoring interest, the daily mark-to-markets lead to a profit of $\$90.07 - \$84.20 = \$5.87$ per barrel on the futures contracts. The company can buy a barrel of oil in the spot market for \$90.10. The total cost of a barrel of oil is therefore $\$90.10 - \$5.87 = \$84.23$ per barrel. This is close to the futures price on 16 March, when the contract was entered.

□

Hedging with futures will be discussed in greater detail in Section3.3.

3.1.4 SAFEX

In South Africa, futures are traded on the South African Futures Exchange (SAFEX), a subsidiary of the JSE. All trading is done via an automated trading system.

You can find out more about SAFEX at http://www.safex.co.za. The following table contains some of the history of SAFEX, taken from this site.

Date	Event
October 2009	Exchange starts valuing stock options off volatility skew.
May 2009	Introduction of Pre– and Post–Trade Anonimity on all equity derivatives.
November 2008	International Derivatives (IDX) allows South African investors gain exposure international shares and indices through the JSE.
June 2006	Can–Do's start trading.
May 2001	Safex and JSE members agree to buyout of Safex by JSE Securities Exchange. Effective date of transaction to be 1 July 2001. The JSE agrees to retain the Safex branding and creates two divisions - Safex Financial Derivatives and Safex Agricultural Derivatives.
August 2000	Individual Equities contract listed has increased to 49.
July 2000	New Government Bond Index launched (GOVI).
February 1999	The Individual Equity Options are replaced with twelve Individual Equity Futures and Options on the futures. Twelve IEF's listed.
March 1998	Options introduced on agricultural products.
September 1997	Individual equity options introduced on the six largest equity counters.
January 1997	Open interest exceeds 1 million contracts.
May 1996	Introduction of fully-automated trading through a specifically designed system that was written in South Africa.
January 1995	Safex Agricultural Derivatives Division opened.
December 1993	Volumes exceed 1 million per month for the first time. Open interest is over 500, 000 contracts.
January 1993	Monthly volumes exceed 200,000 contracts. Open interest exceeds 100,000 contracts.
October 1992	Options-on-futures launched together with a world- class portfolio scanning-ype margining system.
June 1992	Monthly volumes start consistently exceeding 100,000 contracts.
October 1991	Permission received from the South African Reserve Bank for non-residents to participate on Safex via the Financial Rand system.
August 1990	Enabling legislation (the Financial Markets Control Act 1990) is enacted and Safex is officially licensed as a derivatives exchange. Officially opened on 10 August 1990 by the Minister of Finance. Monthly volumes are approximately 60,000 contracts with 10,000 open interest.
April 1990	Safcom takes over operation of the informal futures market from RMB. Futures contracts are available on equity indices, long bonds and money market products.
September 1988	Twenty-one banks and financial institutions meet and establish the South African Futures Exchange (Safex) and the Safex Clearing Company (Safcom).
April 1987	Rand Merchant Bank Limited (RMB) start 5 trading "futures"contracts on various equity indices and long bonds. RMB is the exchange, clearing house and only market maker.

History of SAFEX.
Source: www.safex.co.za

FTSE/JSE Top 40 Index Future

Futures Contract Code	Underlying Instrument	Contract Size	Expiry Dates & Times	Quotations	Minimum Price Movement	Daily Valuation Method	Expiry Valuation Method	Settlement Method
ALSI	FTSE/JSE Top 40 Index (J200)	R10 per point	An intraday auction that starts at noon 12h00 SA time on 3rd Thursday of Mar, Jun, Sep & Dec. (or previous business day if a public holiday)	In whole index points	One Index Point (R10) on screen Off screen trades can be booked to 4 decimal places	Random Snapshot using the middle of the double between 17h00 and 17H05 For further dated expiries a fair value process is used	The intraday Auction is based on a volume maximizing algorithm between 12h00 and 12h15, if there are large price movements in a constituent stock its auction will be extended by 4 minutes. If the price remains far apart from its last pre auction trade its auction will be extended a further 4 minutes. There could be a 2 minute extension independent of the other extensions if there is an imbalance between market orders and limit orders. Thus the maximum length of an auction can be 25 minutes The uncrossing price from this auction will be the expiry price	Cash Settled

Figure 3.2: The ALSI futures contract specifications. Source: SAFEX

3.2 Arbitrage Arguments for Forward Contracts

In this section, we introduce the notion of arbitrage, and use it to determine the theoretical forward price of a given forward contract. Later, we shall see how arbitrage can be used to price a variety of financial derivatives.

3.2.1 Assumptions for Arbitrage Pricing

The academic definition of arbitrage is as folows:

Definition 3.2.1 A *arbitrage opportunity* is a trading strategy that costs nothing to enter, has a non–zero probability of making a profit, and a zero probability of making a loss.

□

We make the following assumptions:

1. **Frictionless markets:** This means
   (i) No transaction costs;
   (ii) No taxes (or else the same taxation rate for all investors);
   (iii) No bid–ask spreads;
   (iv) No margin requirements;
   (v) No restrictions on short sales (or else the existence of a large number of investors who hold the asset for investment purposes);
   (vi) Unlimited borrowing and lending at the same risk–free rate.
2. **Absence of credit risk:** There is no chance of investors defaulting on any obligations.
3. **Liquid and competitive markets:** Investors can buy/sell an asset in any quantity without affecting the price, i.e. they are *price takers.*
4. **Arbitrage–free markets:** Investors snap up any arbitrage opportunity the moment it occurs.

An arbitrage opportunity is a highly desirable thing. It costs nothing, will guarantee that no losses can occur, and may well give you a profit. Think of an arbitrage opportunity as a *free lottery ticket.* If you were offered free lottery tickets, how many would you take? Why, *all* of them, of course! That's why usually there aren't any free lottery tickets to be had. The assumption that there are no arbitarge opportunities is fundamental for financial mathematics.

Consider the following: Corporation ABC, like most other listed companies, has issued two kinds of shares. Some are even–numbered and others are odd–numbered. In all other respects, they are the same, i.e. they confer the same voting rights, and will pay the same dividends. Why should the even–numbered shares have the same price as the odd–numbered shares? The answer is arbitrage. Suppose, for example, that even–numbered shares cost $20.00 and that odd–numbered shares cost $21.00. The astute investor will short an odd–numbered share and use the proceeds to buy an even–numbered share. The difference of $1.00 is free money, because all the obligations due to the short odd–numbered share (i.e. dividends) can be met by the proceeds of the long even–numbered share. Investors will therefore short odd–numbered shares in droves (causing the price to drop), and use the proceeds to buy even–numbered shares (causing their price to rise). This will stop once even–numbered and odd–numbered shares cost the same.

The idea is that arbitrage acts as a kind of force which keeps prices in line. The moment an arbitrage opportunity arises, so many investors will rush to take advantage of the opportunity of a free lunch that the laws of supply and demand will cause prices to readjust until the arbitrage opportunity disappears. Note that this conflicts with Assumption 3! However, we only require these assumptions to be approximately true, or true for a number of key market participants.

3.2.2 Forward Prices

In this section, we consider forward prices on equity, indices and foreign exchange. We shall show that these forward prices are determined not by supply and demand, but by arbitrage arguments.

An arbitrage argument for the forward price

Consider a forward contract that obliges us to hand over an amount F at time T in return for an asset S. The current date is t and the current asset (spot) price is S_t. We don't know what the price S_T of the asset will be at time T, so we cannot tell whether or not we will make a profit. Nevertheless, no matter what we think the price movement will be, we will agree on the forward (delivery) price F. Before we see why, we introduce some terminology:

Suppose that the risk–free interest rate is r. We will assume that interest is continuously compounded, so that an amount P_0 deposited at time t will grow to $P_0 e^{r(T-t)}$ at time T. Consider now the following portfolio:

Action	Cashflow at time t	Cashflow at time T
Long forward	0	$S_T - F$
Short asset	S_t	$-S_T$
Lend cash at r	$-S_t$	$S_t e^{r(T-t)}$
Total	0	$S_t e^{r(T-t)} - F$

The main thing to notice is that the final time–T value of the portfolio is already known at the initial time t, i.e. there is *no risk* involved in holding on to this portfolio. Since the portfolio has zero initial value, it should also have zero final value, i.e.

$$\boxed{F(t,T) = S_t e^{r(T-t)}}$$

is the forward price at time t for the delivery of the asset S at time T.

You may not be convinced at this point, so we'll look at it in another way. Suppose that the forward price F is greater than $S_t e^{r(T-t)}$. In that case you can make a risk–free profit as follows: At time t, enter the short side of a forward contract, borrow S_t from the bank and use this money to buy the asset. Your net cashflow is $0 + S_t - S_t = 0$. At time T you have the following: One forward contract obliging you to deliver the asset S_T for a price F, one asset S_T (which you will use to honour your short position in the forward contract) and a debt of $S_t e^{r(T-t)}$ owed to the bank. Since $F > S_t e^{r(T-t)}$ you have made a net profit from an initial investment of nothing! This is called an **arbitrage opportunity**, and is very desirable. It cannot exist for long, because numerous investors will do the same: They will all short a forward contract, borrow S_t from the bank, and buy the asset. The increased demand for the asset will increase its price, until the arbitrage opportunity disappears, i.e. when $F(t,T) = S_t e^{r(T-t)}$. Arbitrage is the force which keeps prices in equilibrium.

Similarly, when $F(t,T) < S_t e^{r(T-t)}$ you can take advantage of an arbitrage opportunity by entering on the long side of a forward contract, short selling the asset, and putting the resulting cash (i.e. S_t) in a risk–free bank account. Try it!

We thus have $F(t,T)e^{-r(T-t)} = S_t$, i.e.

the discounted forward price = the current spot price .

Note also that $F(T,T) = S_T$, as one would expect.

What if the underlying share pays dividends? To find out, do the following exercise:

Exercise 3.2.2 (a) Consider a forward contract on a dividend–paying asset S. The contract is entered into at time t and initially has zero value. At time T (expiry) the holder is obliged to buy the asset for an amount F (forward price). If it is known at time t that

the underlying asset will pay a dividend of D at time t_d, where $t \le t_d \le T$, calculate the forward price using an arbitrage argument. Assume that the risk–free rate is a constant r.

(b) Consider the same situation, except that this time the dividend paid at t_d is a known fraction of the share price, i.e. the dividend is DS_{t_d}, where $0 \le D \le 1$. The actual dividend is unknown at time t, because the share price S_{t_d} is unknown. Again calculate the forward price by an arbitrage method.
[Hint: Take the view of the writer (short position), and reinvest the dividend in the asset the moment it is paid out.]

Answers: (a) $F(t,T) = S_t e^{r(T-t)} - De^{r(T-t_d)}$; (b) $F(t,T) = \frac{S_t e^{r(T-t)}}{1+D}$.

□

Forward price of an asset with known cashflows

In the above, we have assumed that the interest rate is constant for all maturities. In reality, this is not the case — The rate for a one–year deposit may not be the same as the rate for a two–year deposit. Nevertheless, we can still calculate forward prices using the correct riskless rate implied by a zero coupon bond of the same maturity as the forward contract. Let $B(t,T)$ be the time–t value of a riskless zero coupon bond that pay out \$1.00 at time T. Let S be an asset, and suppose that C is the *present value* of all cashflows associated with the asset over the life of a forward contract initiated at time $t = 0$, with maturity $t = T$. Then the forward price $F(0,T)$ of the asset is

$$F(0,T) = \frac{S_0 - C}{B(0,T)}$$

To see why, consider the following argument:

- At $t = 0$ long one forward contract (which costs nothing) and short one share (which provides income of S_0). Use this to buy $\frac{S_0}{B(0,T)}$ zero coupon bonds (which costs S_0). The initial value of this portfolio is \$0.00.

- At $t = T$, you must pay $F(0,T)$ to receive one share. You will also receive $\frac{S_0}{B(0,T)}$ from the zero coupon bonds. Thus in total you receive $S_T + \frac{S_0}{B(0,T)}$.
  The share must be returned to the party from whom it was borrowed. You must also return all dividends that have been paid in the interim. The $t = 0$–value of the dividends is C, so the $t = T$–value is $\frac{C}{B(0,T)}$. Thus you must pay a total $F(0,T) + S_T + \frac{C}{B(0,T)}$.

- The net cashflow at $t = T$ is therefore

$$S_T + \frac{S_0}{B(0,T)} - F(0,T) - S_T - \frac{C}{B(0,T)} = \frac{S_0 - C}{B(0,T)} - F(0,T)$$

  This future cash flow is known at time $t = 0$: The only source of randomness, S_T, has been cancelled out.

- Since the portfolio cost nothing initially and has a known payoff, its known payoff must be zero, or else there will be arbitrage. Thus $\frac{S_0 - C}{B(0,T)} - F(0,T) = 0$.

Perhaps the easiest way to remember this is: $F(0,T)B(0,T) = S_0 - C$, i.e.

$$\text{Discounted forward price} = \text{Spot price} - \text{PV(Cashflows)}$$

Forward prices an assets with a known dividend yield

Sometimes the actual magnitude of a future dividend is unknown. Instead, it is estimated that future dividends will be some known proportion of the asset price. To simplify, as modelling assumption, we assume that the dividend is paid out in continuous instalments, like continuously compounded interest. This is not a bad assumption for certain securities, e.g. stock indices.

When we say that an asset S has a continuous dividend yield of q, we mean that it pays "interest" at an annualized continuously compounded rate of q on the value of the asset: It pays, between times t and $t+\Delta t$, and amount of approximately $qS_t\Delta t$. Thus future dividends are random, because S_t is unknown. To get around this, note that with $qS_t\Delta t$ it is possible to buy exactly $q\Delta t$ shares. Thus we can remove some of the randomness by reinvesting the dividend payouts in the underlying share.

Suppose that we start off with one share at $t = 0$. At time Δt, we will receive a dividend of $qS_t\Delta t$. We immediately use this to buy $q\Delta t$ shares. we thus have $1 + q\Delta t$ shares at time Δt.

Suppose that at time t we have n_t shares. Then at time $t + \Delta t$, we will receive $n_t qS_t\Delta t$ in dividends. If this is reinvested in the share, we will own $n_t(1 + q)\Delta t$ shares. It follows that

$$dn = nq\ dt$$

Solving this DE, we see that

$$n_t = n_0 e^{qt}$$

Thus the number of shares we own at any time is deterministic, though the share price itself is random.

We can now use an arbitrage argument to calculate the forward price of an asset with a known dividend yield:

Action	Cashflow at time $t = 0$	Cashflow at time $t = T$
Long 1 share Reinvest all dividends in the share	$-S_0$	$S_T e^{qT}$
Short e^{qT} forwards	0	$(F(0,T) - S_T)e^{qT}$
Borrow S_t at a rate of r	S_0	$-S_0e^{rT}$
Total	0	$S_0e^{rT} - F(0,T)e^{qT}$

It follows that

$$F(0,T) = S_0e^{(r-q)T}$$

If interest rates are not constant, then

$$F(0,T)B(0,T) = S_0e^{-qT} \qquad F(t,T)B(t,T) = S_te^{-q(T-t)}$$

Thus the *discounted forward price* of one share is the *current value of the number of shares* you must own now to ensure that you possess one share at maturity.

Forward contracts on a stock index are usually modelled by assuming that the index pays a continuous dividend yield.

Forward contracts on foreign currency

Different countries have different risk–free rates. This difference in interest rates is a major deterministic reason for why currency exchange rates are not fixed: If the risk–free rate in the UK is 6%, and in the USA is 2%, we expect the \$/£–rate to decrease. In one year's time, the dollar cost of a pound is expected to be less than what it is today. See Example **??** for an intuitive explanation of this. The aim of this section is to make this reasoning precise.

Suppose that we want to enter into a forward contract which entitles us to buy one unit of foreign currency at time T. What is the correct forward price (i.e. exchange rate) which sets the initial value of this contract to zero? Let local currency be denoted \$, and foreign currency £. Let S_t denote the spot \$/£–exchange rate at time t, i.e. the time–t value of £1 in \$. Let r_l and r_f, denote, respectively, the local and foreign risk–free rates over the period $[0, T]$. Also let the forward price of £1 be \$$K$. We must find K.

Action	Cashflow at time $t = 0$	Cashflow at time $t = T$
Borrow £ 1	£$1 = \$\ S_0$	$-£\ e^{r_f T} = -\$\ S_T e^{r_f T}$
Lend \$ S_0 to local bank	$-\$\ S_0$	$\$\ S_0 e^{r_l T}$
Long $e^{r_f T}$ forwards	0	$[£\ 1 - \$\ K]e^{r_f T} = \$\ [S_T - K]e^{r_f T}$
Total	0	$\$\ S_0 e^{r_l T} - K e^{r_f T}$

The cashflow at $t = 0$ is zero, and the cashflow at $t = T$ is already known at $t = 0$ (i.e. deterministic). It follows that

$$K = S_0 e^{(r_l - r_f)T}$$

or else there is arbitrage.

There is a simpler way to see this. Cash can be considered as a security that pays a continuous dividend yield, namely the risk–free rate. Thus £ 1 is a "security" that pays a continuous dividend yield equal to r_f. The forward price of an asset S paying a continuous dividend yield q has already been calculated above to be $K = S_0 e^{(r-q)T}$. Since $r = r_l$ and $q = r_f$, we get obtain $K = S_0 e^{(r_l - r_f)T}$ once more.

Forward contracts on commodities

Commodities are physical goods. They therefore may have significant storage costs. For example, the party who is short a forward or future on gold is obliged to deliver a certain quantity of gold at a future date. This gold has to be stored safely until the delivery date, and storage facilities for gold are not free. Hence the party who is long a gold forward contract will be charged for storage. This is most easily viewed as a negative dividend or cashflow. Thus

$$F(0, T) = (S_0 + C)e^{rT}$$

where S_0 is the current spot price, and C is the present value of the storage costs over the period $[0, T]$.

We chose gold, rather than grain, in the above example, because gold is generally kept as an investment asset. The argument above breaks down for consumption commodities, such as grain, for the simple reason that grain is not kept as an investment asset: Companies who keep grain in inventory do so because they need to "consume" it as part of their production process. To see why this affects the forward price, we must briefly consider the arbitrage argument leading to the forward price.

Let S_0 be the spot price of a bushel of grain, F its forward price, and let C be the present value of storage costs. Suppose that

$$F > (S_0 + C)e^{rT}$$

In that case, a flour factory

- Borrow $S_0 + C$ and use it to buy one bushel of grain, and to pay for storage.
- Short a forward on one bushel of grain.

The flour factory needs to have grain in order to function, and is therefore happy to buy large amounts of grain. Hence the spot and forward prices will adjust until F is no longer greater than $(S_0 + C)e^{rT}$.

Now suppose that

$$F < (S_0 + C)e^{rT}$$

To make use of the arbitrage opportunity, the flour factory must

- Sell one bushel of grain, save the storage costs, and lend $S_0 + C$ at the risk–free rate.
- Long a forward on one bushel of grain.

But the factory will not generally be prepared to sell grain and buy forwards: After all, grain is needed for production on a day–to–day basis, and forward contracts can't be milled and than sold! Hence the force of arbitrage which pushes F up to $(S_0 + C)e^{rT}$ is not very strong. For consumption commodities, arbitrage works only one way, i.e.

$$F(0, T) \leq (S_0 + C)e^{rT}$$

In the above, we only considered storage costs. However, some commodities also bestow a benefit on the owner. For example, a flour factory would much rather have grain than forward contracts. Ownership of grain enables the factory to keep its production going.

The benefit of holding a physical commodity is called a *convenience yield*, and can be viewed as a positive dividend or cashflow.

3.2.3 Valuation of Forward Contracts

The forward price of an asset S is calculated so that the *initial value* of a forward contract (with delivery price equal to the forward price) is zero (or there would be arbitrage). However, this is only true initially. Suppose that the agreed–upon delivery price is K. At maturity T, the value of the forward contract is $S_T - K$ to the party who is long, and $K - S_T$ to the party who is short, and these numbers generally are not zero.

Over time, the asset price changes, and so does the forward price. However the delivery price agreed upon in the forward contract does not change. If the forward price becomes

greater than the delivery price, then the party who is long a forward contract has to pay less than the market–agreed fair price, and thus the forward contract has positive value for the party who is long. Similarly, if the forward price moves below the delivery price, the contract has negative value to "long", and positive value to "short".

The valuation of forward contracts is a trivial matter. Let S_t and $F(t,T)$ denote, respectively, the spot and forward prices of an asset S, where the forward price is for maturity T. Let K be the delivery price of a forward contract initiated at $t = 0$, i.e. $K = F(0,T)$. At a later time t, the forward price of the asset has changed to $F(t,T)$, i.e. the fair price of S for delivery at time T is $F(t,T)$. However, the party who is long the original forward contract must pay K, and not $F(t,T)$. This is a profit (or loss) of $F(t,T) - K$ at time T, and thus $(F(t,T) - K)e^{-r(T-t)}$ at time t. Hence the time–t value of the forward contract is $(F(t,T) - K)e^{-r(T-t)}$ to "long".

To convince you that this is the case, consider the following argument. Suppose that at $t = 0$ you went long one forward contract with maturity T and delivery price K. At time t, you can lock in a profit (or loss) of $(F(t,T) - K)e^{-r(T-t)}$ as follows:

- At time t short one forward contract with maturity T and delivery price $F(t,T)$.
- At time T the contract you are long obliges you to pay K in return for S_T.
- The contract you are short requires you to pay S_T in return for $F(t,T)$.
- Net cashflow at time T is therefore $F(t,T) - K$. This is known at time t.
- The time–t value of this is $(F(t,T) - K)e^{-r(T-t)}$.

If interest rates are not constant, it is easy to see that the time t–value of a forward contract is just $(F(t,T) - K)B(t,T)$ to the party who is long.

Note that, since $K = F(0,T)$, the $t = 0$–value of a forward contract, as determined by the above formula, is indeed zero.

3.2.4 Futures Prices

Futures contracts are similar to forward contracts, in that both are obligations to pay a certain price for an asset at a future date. Nevertheless, the differences in settlement methodologies (day–to–day marking to market vs. settlement at maturity) introduce important differences in future and forward prices. This uncertainty is due entirely to the fact that interest rates fluctuate randomly.

Marking–to–market involves the daily settlement of profits and losses by paying into a margin account. Amounts deposited in the margin account will offer the risk–free rate of interest, but as this is randomly changing from day to day, there will be some uncertainty in the final payoff due to interest rates (in addition to uncertainty due to spot price fluctuations).

Let $F(t,T)$ and $\mathcal{F}(t,T)$ denote, respectively, the time–t forward and futures settlement prices of an asset S for delivery at T. A long futures contract will have the following sequence of cashflows into the margin account (other than margin calls) associated with it:

Day	Cashflow
0	0
1	$\mathcal{F}(1,T)-\mathcal{F}(0,T)$
2	$\mathcal{F}(2,T)-\mathcal{F}(1,T)$
3	$\mathcal{F}(3,T)-\mathcal{F}(2,T)$
$\vdots$	$\vdots$
T	$S_T-\mathcal{F}(T-1,T)$

Now let i_n denote the simple daily[2] interest rate prevailing form the end of day n to the end of day $n+1$. The first cashflow $\mathcal{F}(1,T)-\mathcal{F}(0,T)$, deposited at the end of the first day, will grow to

$$[\mathcal{F}(1,T)-\mathcal{F}(0,T)](1+i_1)(1+i_2)\dots(1+i_{T-1})$$

The second cashflow $\mathcal{F}(1,T)-\mathcal{F}(0,T)$, deposited at the end of the second day, will grow to

$$[\mathcal{F}(2,T)-\mathcal{F}(1,T)](1+i_2)(1+i_3)\dots(1+i_{T-1})$$

The total payoff at maturity is therefore

$$\sum_{n=1}^{T}\left([\mathcal{F}(n,T)-\mathcal{F}(n-1,T)]\prod_{k=n}^{T-1}(1+i_k)\right)$$

and this is random if the $i_1,\dots,i_{T-1}$ are unknown.

Suppose, however, that interest rates are deterministic, i.e. that at $t=0$ it is already know what $i_1,\dots,i_{T-1}$ will be. We then claim that the forward and the future price are the same, i.e. that $F(0,T)=\mathcal{F}(0,T)$. We will show this using an arbitrage argument.

Let $m_k=(1+i_k)$ be the inflator over the k^{th} day, and let $\Delta\mathcal{F}_k=\mathcal{F}(k,T)-\mathcal{F}(k-1,T)$ be the change in futures price over the k^{th} day. Consider the following trading strategy, which we shall call Portfolio A:

Day	Action at end of day	Settlement at end of *next* day	Future Value of settlement
0	Long $m_1=(1+i_1)$ futures	$m_1\Delta\mathcal{F}_1$	$m_1m_2\dots m_{T-1}\Delta\mathcal{F}_1$
1	Long an additional i_2m_1 futures Total: $m_1(1+i_2)=m_1m_2$ futures	$m_1m_2\Delta\mathcal{F}_2$	$m_1m_2\dots m_{T-1}\Delta\mathcal{F}_1$
2	Long an additional $i_3m_1m_2$ futures Total: $m_1m_2(1+i_3)=m_1m_2m_3$ futures	$m_1m_2m_3\Delta\mathcal{F}_3$	$m_1m_2\dots m_{T-1}\Delta\mathcal{F}_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$T-1$	Long $i_{T-1}m_1\dots m_{T-2}$ futures Total: $m_1m_2\dots m_{T-1}$ futures	$m_1\dots m_{T-1}\Delta\mathcal{F}_T$	$m_1\dots m_{T-1}\Delta\mathcal{F}_T$

Thus the total value at maturity is $m_1\dots m_{T-1}[\Delta\mathcal{F}_1+\dots+\Delta\mathcal{F}_T]$ (the sum of the right–hand column). But clearly $\Delta\mathcal{F}_1+\dots+\Delta\mathcal{F}_T=S_T-\mathcal{F}(0,T)$. Thus **Portfolio A** has a payoff of

$$m_1m_2\dots m_{T-1}(S_T-\mathcal{F}(0,T))$$

at maturity.

Now consider the following trading strategy, Porfolio B:

[2]i.e. *not* annualized.

- **Portfolio B:** At end of day 0, short $m_1 \dots m_T$ forward contracts with forward price $F(0,T)$.

We can do this at $t = 0$, because $m_1, \dots, m_{T-1}$ are *known*. Thus if interest rates are deterministic, Portfolio B has a payoff of

$$-m_1 m_2 \dots m_{T-1}(S_T - F(0,T))$$

at maturity.

Now consider the combined portfolio: Portfolio C = Portfolio A + Portfolio B. Portfolio C has an initial cost of zero, since neither forwards nor futures cost anything. The payoff of **Portfolio C** is

$$m_1 \dots m_{T-1}[F(0,T) - \mathcal{F}(0,T)]$$

at maturity, but this is already known at time $t = 0$. Thus

$$\mathcal{F}(0,T) = F(0,T) \quad \text{if interest rates are deterministic}$$

or else there will be arbitrage.

We have therefore shown that if interest rates are deterministic, futures and forward prices coincide. In particular, this is true if we make the assumption that interest rates are constant (as we will often do).

If interest rates are random, however, futures and forward prices are generally not the same. Suppose, for example, that the underlying asset is positively *positively correlated* with interest rates. Then:

Rates increase	$\Longrightarrow$	Underlying asset price increases
	$\Longrightarrow$	Futures price increases
	$\Longrightarrow$	"Long" receives a positive settlement which can be invested at higher rates.
Similarly,		
Rates decrease	$\Longrightarrow$	Underlying asset price decreases
	$\Longrightarrow$	Futures price decreases
	$\Longrightarrow$	"Long" makes a positive payment which can be borrowed at lower rates.

Thus whether interest rates increase or decrease, "long" receives a positive benefit from rate changes: If rates increase, "long" will receive cash, to be invested at high rates. If rates decrease, "long" must pay, but can borrow this at low rates.

Thus if interest rates are positively correlated with the underlying asset, the party who is long a futures contract has an advantage over someone who is short. This advantage must be paid for, and "long" will be required to make a higher payment at delivery. Thus, all things being equal, the futures price of an asset tends to be greater than the forward price if interest rates are positively correlated with the underlying.

The opposite is true if interest rates are negatively correlated with the underlying asset. Examples of such assets are bonds, whose values decrease if interest rates increase. Thus the futures price of a bond tends to be lower than the forward price.

We end this section by briefly discussing some other factors which will move futures prices away from forward prices. One argument is that the futures price of an asset is the *expected price of the asset* at delivery time. Immediately, one can ask: Expected by whom?

Expectation with respect to what probability measure? The intuitive idea, however, is that investors who go long on futures expect the future spot price to move above the futures price, and that investors who go short expect the future spot price to move below the futures price. Now every futures contract has a party who is long *and* a party who is short. If these are both right roughly half the time, one would expect the future spot price to be roughly equal to the futures price.

Another argument, due to Keynes and Hicks, attacks this line of reasoning. Futures are used for different purposes. It may be that hedgers tend to hold short positions, whereas speculators tend to hold long positions, for example. Hedgers are more concerned with getting insurance, for which they are prepared to pay. Speculators, on the other hand, demand compensation for the risks that they take. Thus if hedgers are short and speculators are long, the expected future spot price will be above the futures price, so that the payoff to the speculators

$$\text{Payoff} \approx S_T - \mathcal{F}(0,T)$$

tends to be positive. The negative payoff to the hedger is just an insurance premium. In this case, the market is said to be in *backwardation*.

Similarly, if hedgers tend to be long and speculators short, then the expected future spot price will be below the futures price. In this case, the market is said to be *contango*.

Futures prices may also move away from forward prices because of certain option–like qualities implicit in a futures contract. Futures are standardized, but the standards allow for a certain flexibility in delivery grade and delivery time. Consider, for example, CBOT–traded T–bond futures, discussed in greater detail in Section 6.3.1. Trading in T–bond futures contracts ceases at 14h00, two hours before trading in the T–bonds themselves. A party who is short a T–bond future can give notice of intention to delivery before 20h00 on any day of the delivery month. Thus if between 14h00 and 16h00 T–bond prices drop drastically, the party who is short can send the clearing house notice of intention to deliver. She may the deliver a bond which is worth only the low 16h00 price, but will receive payment equal to the high 14h00 futures settlement price. Thus the party who is short a T–bond future has an option to make a so–called *wild card play*. Options aren't free, however, and this option will be reflected in a lower T–bond futures price.

Similarly, a T–bond futures contract doesn't specify exactly which bond must be delivered; instead, a variety of T–bond futures can be delivered: A T–bond futures contract specifies that any T–bond with maturity greater than 15 years (and not callable within 15 years) can be delivered. This means that a party who is short has an option to deliver the cheapest such bond. Again, this option is paid for by lowering the futures price.

3.3 Hedging with Futures

We begin this section with two examples, illustrating the use of futures for hedging:

3.3.1 Some Examples

Both the following examples were taken from the *Financial Risk Manager's Handbook 2001–2002*, by P. Jorion, Wiley 2001.

Example 3.3.1
- In the early 1990's, Metallgesellschaft, a German oil company suffered a loss of \$1.33 billion in their hedging programme.

- They sold long-term fixed-price contracts for heating oil and gasoline to their customers, thus creating a long-term exposure to oil prices.
- They hedged their exposure by rolling over short-dated oil futures, but eventually had to abandon their hedge because of large negative cashflows.
- Their hedging team maintained that such cashflows would eventually be balanced out through gain on the long-term fixed-price contracts, but senior management closed out the hedged position because of concern about the large cash drain.)

Question: Which of the following alternatives happened?

(a) MG hedged its exposure by shorting futures, and there was a decline in the oil price;

(b) MG hedged its exposure by longing futures, and there was a decline in the oil price;

(c) MG hedged its exposure by shorting futures, and there was a rise in the oil price;

(d) MG hedged its exposure by longing futures, and there was a rise in the oil price;

Solution:

- MG sold oil at fixed prices, and would therefore suffer if oil prices rose. To hedge against a rise in oil prices, they would have gone long oil futures. (Going short would have increased their exposure to increasing prices.)
- Margin calls on these futures were made when oil prices dropped, causing negative cashflows.
- (Note that dropping prices initially didn't cause any concern to MG, i.e. the hedge caused the problem.)
- Thus: (b) MG was long on futures, and oil prices declined.
- In the long run, the fact that oil prices declined meant that MG would have been able to sell oil to its fixed-price customers for a profit, i.e. the negative cashflows would indeed have been balanced by positive cashflows in the future. (However, there is some disagreement on this. You can find a case study which includes several views in the *Derivatives Handbook*, edited by R.J. Schwartz and C.W. Smith, John Wiley & Sons, 1997.)

□

Example 3.3.2 • A US exporter has been promised a payment of ¥125m in 7 months.

- The perfect hedge would be an OTC forward, but such a contract is not very liquid.
- Instead the exporter opts to use exchange-traded futures contracts.
- The CME lists Yen contracts with face value ¥12.5m that expire in 9 months.

- The exporter shorts 10 futures contracts, with the intention of closing out the position after 7 months, when the futures contracts still have 2 months to maturity. This is an example of a *short hedge.*

Item	**Initial Time**	**Exit Time**	**Profit/Loss**
Market Data:			
Maturity	9 months	2 months	
\$ interest rate	6%	6%	
¥interest rate	5%	3%	
Spot ¥/\$	125.00	150.00	
Futures ¥/\$	124.07	149.00	
Contract Data:			
Spot \$/¥	0.008000	0.006667	-\$166 667
Futures \$/¥	0.008060	0.006711	+\$168 621
Basis \$/¥	-0.000060	-0.000045	+\$1 954

- Over the hedging period, the Yen depreciates sharply, leading to a loss on the anticipated dollar position of

$$\$125000000 \times (0.006667 - 0.008000) = -\$166\ 667$$

  This is the loss that would have been realized without the hedge in place.

- However the loss on the cash is offset by a gain on the futures position:

$$(-10) \times \$12500000 \times (0.006711 - 0.008060) = +\$168\ 621$$

- In fact, the US exporter made a small gain of \$1 954.

- This small gain was unpredictable, and could also have been a small loss. The hedge is therefore not perfect, and the cause of this imperfection is termed *basis risk.*

□

3.3.2 Basis Risk

The *basis* is defined to be the spot price minus the futures price:

$$b_t = S_t - \mathcal{F}(t, T)$$

At maturity T, the basis is zero. Before that time, however, it can vary randomly, and can be both positive and negative. If the spot price increases by more than th futures price, the basis is *strengthened* When the futures price increases by more than the spot price, the basis is *weakened.*

Basis risk arises when the characteristics of the underlying position differ from those of the hedging (futures) position, e.g.

- Maturities may be different so that futures prices have not yet converged to the spot (e.g. above example).

- Futures are standardized: There may not be a position in futures contracts available with exactly the same quantity of the underlying as needs to be hedged (e.g. if exporter was expecting ¥130M)

- The asset underlying the futures contract may be different from the asset that needs to be hedged. They may merely be highly correlated. For example, one can hedge jet fuel using heating oil futures, or Norwegian Kroner using Euro futures. This is called *cross-hedging*, and substantially affects basis risk

In order to use futures to hedge, one must choose:

- Asset that underlies futures contract.
(If same asset is not available, choose asset that is strongly correlated.)

- Delivery month of futures contract.
(Because prices are frequently quite erratic during delivery months, the delivery month of the future may be later than the delivery month of the asset).

Sometimes the delivery date of the asset is greater than the delivery date of any traded futures contract. The hedge can then be *rolled forward*. This involves entering a futures position, closing it out before maturity, and entering a new futures position with a later maturity, closing that out, and entering a new futures position, ... etc. Every time a position is closed and a new one entered, there is a new source of basis risk.

Since short-dated futures contracts are more liquid than long-dated ones, hedgers may prefer to use a roll forward strategy even if there are futures with sufficiently long maturity.

Now assume a futures hedge is put in place at t_1 and closed out at timet_2. The futures contract may mature at a date later than t_2. Let

$$\begin{aligned} S_n &= \text{ spot price at time } t_n \\ F_n &= \text{ Futures price at time } t_n \\ b_n &= \text{ basis } = S_n - F_n \text{ at time } t_n \end{aligned}$$

The hedger (short hedge) receives:

- S_2 from selling the asset;

- Margin payments totalling $-(F_2 - F_1)$ (short futures).

Hence total gain/loss at t_2 is (ignoring interest)

$$S_2 - F_2 + F_1 = F_1 + b_2$$

This is not quite non-random: b_2 is a random variable, but usually small.

Basis risk is even greater when cross hedging: Hedge asset S with futures on asset S^*. The gain/loss at t_2 is

$$S_2 - F_2 + F_1 = F_1 + (S_2 - S_2^*) + (S_2^* - F_2)$$

which shows the basis has two sources of risk:

- The risk that the asset to be hedged has a different price from the asset that underlies the futures.

- The risk that the price of the underlying asset differs from the futures price.

3.3.3 The Optimal Hedge Ratio

Thus far, we have considered only **unitary hedges** where the quantity of assets to be hedged equals the quantity of assets underlying the futures contracts. However, this is not always the best method. Define

$$h = \frac{\text{size of position in futures contracts}}{\text{size of exposure}}$$

Thus a unitary hedge has $h = 1$.

For simplicity, assume that a corporation needs to buy *one* asset S at a future date T. Define

- ΔS = change in value of the asset that is to be hedged (over hedging period $[0, T]$);
- ΔF = the change in futures price over hedging period;
- $\sigma_{\Delta S}$ = standard deviation of ΔS;
- $\sigma_{\Delta F}$ = standard deviation of ΔF;
- $\sigma_{\Delta_S, \Delta_F}$ = covariance of $\Delta_S, \Delta F$;
- ρ = correlation coefficient

If a risk manager attempts to hedge the future purchase of *one* asset S (e.g. jet fuel) by going *long* h futures contracts (on heating oil), the change in value of the portfolio is

$$\Delta P = h\Delta F - \Delta S$$

This change in value is a random variable. To minimize the risk, one should minimize the variance of the profit/loss. Now the variance is just

$$\sigma_{\Delta P}^2 = \sigma_{\Delta S}^2 + h^2\sigma_{\Delta_F}^2 - 2h\sigma_{\Delta_S,\Delta_F}$$

To minimize the variance, differentiate with respect to h and set the derivative equal to 0:

$$\frac{\partial \sigma_{\Delta P}^2}{\partial h} = 2h\sigma_{\Delta F}^2 - 2\sigma_{\Delta S,\Delta F}$$

Thus

$$h = \frac{\sigma_{\Delta S,\Delta F}}{\sigma_{\Delta F}^2} = \rho\frac{\sigma_{\Delta S}}{\sigma_{\Delta F}}$$

Hence if a corporation must by one asset S at some future date, it should long $h = \rho\frac{\sigma_{\Delta S}}{\sigma_{\Delta F}}$ futures contracts. We can obtain the $\sigma_{\Delta S}, \sigma_{\Delta F}$ from market observations as follows: The *volatility* σ_S of an asset is the standard deviation of the annualised returns, i.e.

$$\sigma_S = \text{standard deviation of } \frac{\Delta S}{S_0} \quad \text{where } T = 1 \text{ year}$$

Now the standard deviation $\sigma_{\Delta S}$ of the spot price change over a period of length T is just $S_0\sigma_S\sqrt{T}$. The reason for this is that variances of independent increments add up: $\sigma_{\Delta S}$ is the

standard deviation of the annual change in S. Now the change in S over a period T is given by

$$S_T - S_0 = \sum_{t=1}^{T} (S_t - S_{t-1})$$

Hence

$$\sigma^2_{\Delta S} = \mathrm{Var}(S_T - S_0) = \sum_{t=1}^{T} \mathrm{Var}(S_t - S_{t-1})$$

assuming that the random variables $S_t - S_{t-1}$ are mutually independent.

If we assume that the random variables $S_t - S_{t-1}$ are identically distributed, then they will all have the same variance s^2. Thus

$$\sigma^2_{\Delta S} = s^2 T$$

gives us the variance of the price change over a period of time T, where s^2 is the variance of the price change per unit time.

For example, we could measure measure time in units of days. It follows that if s is the standard deviation of daily price changes, then, then the standard deviation of annual price changes is

$$\sigma^2_{\Delta S} = 365 s^2 \quad \text{i.e.} \quad s = \sigma_{\Delta S}\sqrt{\frac{1}{365}}$$

where $T = 1$ year. Note that 1 day $= 1/365$ years.
Actually, traders use a shorter "year", based on the number of trading days. Thus if there are 252 trading days in the year, they would calculate annual standard deviation as $\sigma = s\sqrt{252}$.

Similarly, the standard deviation of the price change over a period of T units of time is just $s\sqrt{T}$, where s is the standard deviation of the price change per unit time. This is true whether we measure time in days, years or seconds.

To get an accurate measure of volatility, it is convenient to use days as units of time. Volatility is generally quoted in annualized terms, however. Thus if years are our unit of time, then we have

$$\text{volatility of } S = \text{standard deviation of annual return}$$

i.e.

$$\begin{aligned}\sigma_S &= \text{standard deviation of } \ \frac{S_1 - S_0}{S_0} \\ &= \frac{\text{standard deviation of } \ S_1 - S_0}{S_0} \\ &= \frac{s}{S_0}\end{aligned}$$

and thus $s = S_0 \sigma_S$. Now the standard deviation of the price change over a period of T years is

$$\sigma_{\Delta S} = s\sqrt{T} = S_0 \sigma_S \sqrt{T}$$

We have therefore found an expression the standard deviation $\sigma_{\Delta S}$ over a period T in terms of the volatility σ_S. The reason for bothering to do this is that volatility (= standard deviation of returns) is what is generally quoted in the market, and not standard deviation

of price change. (Investors don't care much about changes in price; they care about *returns* on investments.)

Similarly $\sigma_{\Delta F} = F_0 \sigma_F \sqrt{T}$.

It follows that

$$h = \rho \frac{\sigma_{\Delta S}}{\sigma_{\Delta F}} = \rho \frac{S_0 \sigma_S}{F_0 \sigma_F}$$

is the optimal hedge ratio.

Example 3.3.3 An airline knows that it needs to buy 10 000 metric tons of jet fuel in three months. It hedges by taking a long position in heating oil futures. The notional of a futures contract is 42 000 gallons. The spot price of jet fuel is \$277/metric ton, and the futures price of heating oil is \$0.6903/gallon. The volatility of jet fuel is 42.34%, and the volatility of the futures price is 37.18%. The correlation between jet fuel and heating oil futures prices is 0.8243. How many futures contracts should the corporation enter into?

Simplify first: If the corporation wanted to buy only one metric ton of fuel, and if the notional for a futures contract was only one gallon, then the number of contracts that should be enter is

$$\rho \frac{S_0 \sigma_S}{F_0 \sigma_F} = 0.8243 \frac{277 \times 0.4234}{0.6903 \times 0.3718} = 376.67$$

However, each futures contract is for 42 000 gallons. Thus only

$$\frac{376.67}{42000} = 0.00897$$

contracts are necessary to hedge the purchase of 1 metric ton of fuel. To hedge 10 000 metric tons, the corporation must therefore long

$$10\ 000 \times 0.00897 = 89.7 \approx 90$$

futures.

□

By reasoning similar to the above, you will easily show that in order to hedge the future purchase of Q_s assets with price per asset s, a corporation should enter a long position of

$$\rho \frac{Q_s s \sigma_s}{Q_f f \sigma_f}$$

futures contracts, where

- Q_f is the notional of the futures contract (i.e. number of units that a single futures contract is for, e.g. 42 000);
- s, f are the spot– and futures prices *per unit*;
- σ_s, σ_f are the volatilities of the spot– and futures prices;
- ρ is the correlation between the spot– and futures price.

Chapter 4

Options

4.1 Equity Option Contracts

Options were introduced in Chapter 2, but we repeat the definition here:

Definition 4.1.1 An *option* gives the holder the right, but not the obligation, to buy or sell an asset — or more generally, to exchange one asset for another.

(a) A *european call option* gives the holder the right to *buy* an asset S (the *underlying*) for an agreed amount K (the *strike price* or *exercise price*) on a specified future date T (*maturity* or *expiry*).

- The party who undertakes to deliver the asset is called the *writer* of the option.
- Because the holder of an option does not have to exercise the option, the payoff to the holder is never negative. The holder would exercise the option at expiry T if and only if the strike K is less than the spot price of the underlying S_T. In that case, he would get a payoff of $S_T - K \geq 0$.
  If the spot price is less than the strike, — i.e. if $S_T < K$ — the holder would discard the option: Why pay K when you can pay $S_T < K$?
- Thus the payoff to the holder is simply $(S_T - K)+ = \max\{S_T - K, 0\}$.
- Since the payoff is never negative, options are not free. So unlike forward contracts, options cost money: You have to pay the writer of an option a *premium* upfront to enter into an option contract. This premium compensates the writer for the risk that $S_T > K$ at maturity.

(b) Similarly, a *european put option* gives the holder the right to *sell* an asset S (the *underlying*) for an agreed amount K (the *strike price* or *exercise price*) on a specified date future T (*maturity* or *expiry*). It has a payoff of $(K - S_T)^+ = \max\{K - S_T, 0\}$.

(c) An *American call (put) option* also confers the right to buy (sell) an underlying asset S for an agreed strike price K. However, this right can be exercised at *any time* before expiry T, and not just at the expiry date.

□

There are many other kinds of equity options, such as barrier options, rainbow options, basket options, asian options, lookback options... In contrast to these *exotic options*, the european/American call/put options we described above are referred to as plain *vanilla options.*

Note that a call option gives the holder the right to *call* for the underlying asset, whereas a put option confers the right to *put* it to someone. A call option is the right to exchange K in return for S, whereas a put option is the right to exchange S in return for K.

An option is said to be *in–the–money* if it has positive payoff, were you able to exercise it straightaway. Thus a call option is in–the–money when $S \geq K$, whereas a put is in–the–money when $S \leq K$. If $S \approx K$, then the option is said to be *at–the –money.* An option which is not in– or at–the–money is said to be *out–of–the–money.* A call (put) which has $S \gg K$ ($S \ll K$) is said to be *deep* in–the–money, etc.

Options can be OTC or exchange–traded (listed). In the latter case, they are standardized, and there are only a few maturities avalaible for each contract. For example, an exchange traded equity call option traded on the CBOE (Chicago Board Options Exchange) is generally of American type, and is for 100 shares of the underlying. Maturities are spaced at three month intervals (e.g. January, April, July, October) and are typically less than one year (though long–term options known as LEAPS have expiry dates up to 39 months ahead, but always expire in January). The exchange will also specify the range of possible strikes, position limits, what happens when dividends are declared, how to adjust for stock splits, etc.

Some exchanges, including SAFEX, will use a mark–to–market approach for traded options that is similar to the one used for futures: losses and gains are calculated on a daily basis and accumulated in a margin account.

4.2 Factors Which Affect Option Prices

The holder of an option will benefit only if the option expires in–the–money. Now future stock prices cannot be predicted accurately. Instead, there are a number of models that purport to describe the stochastic dynamics of the underlying asset S, from which the probability distribution for the future price S_T can be calculated. Given such a model, we expect that the value of the option in that model is determined by:

1. The probability that the option expires in–the–money;

2. The expected payoff, given that the option expires in–the–money.

We isolate the following factors as important determinants of stock option prices:

- Today's stock price S.
- The strike price K.
- The time to expiry T.
- The riskless interest rate r (for the period $[0, T]$).
- The expected dividend yield q.
- The stock price volatility σ.

At the expiry T of an option, only the stock price S and the strike price K determine the payoff. However, these will be important determinants *before* expiry as well. It should be clear that call option prices will be increasing functions of S and decreasing functions of K: If S increases, the option has a greater chance of expiring in–the–money. The opposite is true of K increases.

Similarly, put values are decreasing in S and increasing in K. Indeed, clearly the right to exchange A in return for B becomes more valuable as A increases and B decreases.

The value of the riskless rate r is also a determinant of stock option prices. Broadly speaking, there are two reasons for this:

1. Stock prices are risky assets, so the expected return on a stock is greater than the riskless rate:

$$\text{Expected return} = \text{Riskless rate} + \text{ Risk premium}$$

   Thus if the riskless rate increases, then S_T is expected to increase, and thus call option values will increase, whereas put options will decrease.

2. The holder of a call option has the right to pay K at expiry. Now K at expiry is worth $PV(K)$ now. As the riskless rate increases, the present value of the strike $PV(K)$ decreases. Thus the current value of what is to be paid is less, and thus the current option value is less.

   Similarly, the holder of a put option has the right to receive K at expiry, and that is worth $PV(K)$ now. Hence as the riskless rate increases, the value of a put option decreases.

Thus both the above reasons ensure that call option prices are increasing functions of r, whereas put option prices are decreasing functions of r.

The *volatility* of a stock price is a measure of the uncertainty of the future stock price: The greater a stock's volatility, the more wiggly the graph of the stock price vs time, and the more uncertain the future stock price. Formally, the volatility of a stock is defined to be the *standard deviation of annualized returns.*

For the moment, greater volatility implies greater uncertainty. Thus the volatility is a fundamental determinant of the probability distribution of the future stock price: An increase in volatility will tend to spread out the probability density fucntion of the future stock price, lowering its peak and fattening the tails.

In Figure 4.2, the two stocks shown move together: When one increases or decreases, so does the other. However, one (shown in red) will increase or decrease in much smaller steps than the other, and is said to be *less volatile.*

Heuristically, if the stock price has greater volatility, then the option has a greater chance of expiring deep in–the–money. Thus, *conditional* on the event that the option expires in–the–money, an option on a high volatility asset is likely to have a greater payoff than an option on a low volatility asset (everything else assumed equal.) Of course, the option also has a greater chance of expiring deep out–of–the–money, but that doesn't affect the final payoff: The payoff is zero whether the option expires just out–of–the–money or deep out of–the money. Hence both call and put prices are increasing functions of volatility.

This fact can be easily understood if one sees options as a kind of insurance:

- A call option is insurance against the risk that S rises: No matter how big the value of S at expiry, the holder of a call will pay no more than K.

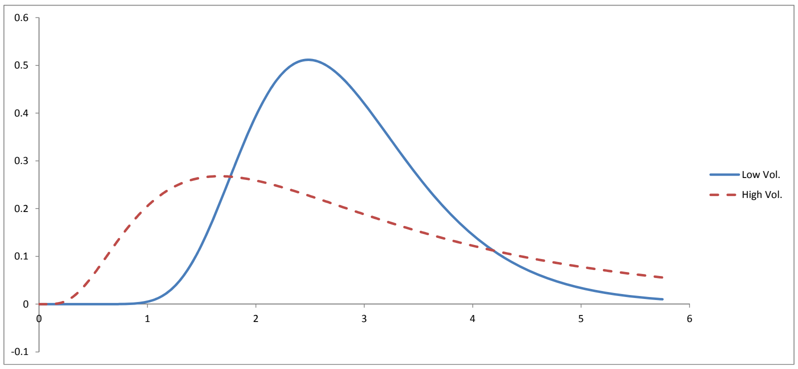

Figure 4.1: Density functions of future stock prices, with the same mean, but different volatilities.

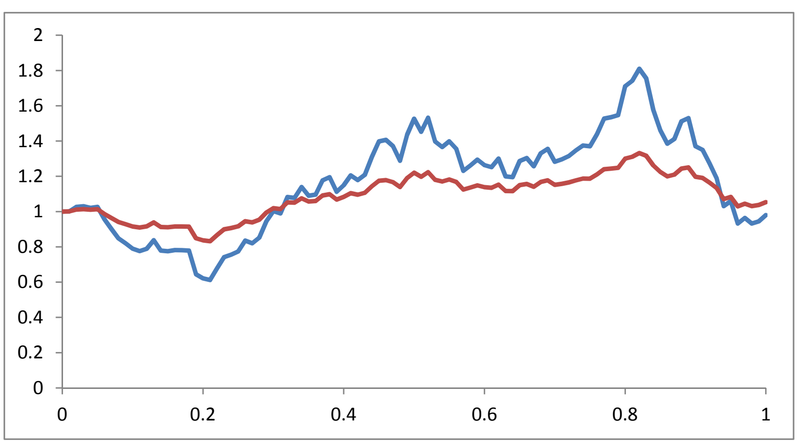

Figure 4.2: Simulated stock prices with high and low volatility.

- Similarly, a put option is insurance against decreasing prices: No matter how small the value of S is at time T, the holder of a put will receive at least K for it.

Obviously, the greater the uncertainty, the greater the insurance premium you have to pay. Thus, the greater the volatility, the greater the option premium.

Similarly, since long–term insurance costs more than short–term insurance, we expect that both call and put prices are increasing functions of the time to expiry T. This is certainly the case for American options. For example if A owns a 2–year american option and B a 1–year American option (both on the same stock with the same strike), then A has greater rights than B does: A can exercise his option any time B does, and at other times besides. Thus for American options it is certainly the case that call and put prices increase with time to expiry. For european options, we have to be more careful. We shall soon see that American calls and european calls on a non–dividend paying stock have the same value, because it is never optimal to exercise an American call on such a stock before expiry[1] . Thus european calls are increasing functions of maturity, because American calls are. For european put options, however, the situation is more complicated. A deep in–the–money european put option may decrease as we increase the time to expiry. [To see this, consider the extreme case of a put with strike K on a worthless stock, i.e. a stock with $S_0 = 0$. Then $S_T = 0$ with probability 1 as well (for if there was a chance that $S_T > 0$ in the future, then we would have $S_0 > 0$ today...) and hence the today's put price is $PV(K) = Ke^{-rT}$ — a decreasing function of T.] For typical values of stocks and strikes, however, european put prices are increasing functions of maturity as well.

Increasing Parameter	Call Price	Put Price
Stock Price	+	−
Strike Price	−	+
Interest Rate	+	−
Time to Expiry	+	?
Volatility	+	+
Dividends	−	+

Effects of parameters on option values.

4.3 Arbitrage Bounds For Option Values

In this section, we use arbitrage arguments to determine upper and lower bounds for option prices. Unlike in the case of forward contracts, we will not yet be able to use arbitrage to determine the exact value of an option contract.

Arbitrage can be used to price derivatives contracts only when it is possible to construct a *replicating portfolio* for the derivative. For example, one can construct a portfolio consisting of stocks and bonds which will exactly replicate the payoff of a forward contract, and arbitrage dictates that the value of this portfolio and that of the forward contract must be the same. (If not, buy the cheaper and short the dearer...) For forward contracts, this replicating portfolio

[1]However, this is not the case if the stock pays dividends before expiry. It may be optimal to exercise an American option just before the stock goes ex–dividend, rather than at expiry. Moreover, if a large dividend is expected in 6 weeks, then a one–month european call my well be more expensive than a two–month european call.

is *static*: You can buy/sell the components of the portfolio now, and do not need to alter or rebalance your portfolio — just let it sit there, and at maturity it will duplicate the payoff of the forward contract exactly.

Options can *also* be replicated with a portfolio of stocks and bonds, and hence *can* be priced. The difference is that this cannot be done statically, but only *dynamically*: the replicating portfolio for an option must be constantly rebalanced. Thus one needs to observe the underlying stock price throughout the life of the option. Hence, in order to price the option, one therefore needs a *stochastic model* of the underlying stock price dynamics — and this will ahve to wait till a later chapter.

For this reason, different models of stock prices will give different prices for the same option. But all models will agree on the value of a forward contract — and all models will agree on the bounds that we establish in this section.

In this section, let

$$\begin{aligned}
S_t &:= \text{price of underlying asset at time } t \\
K &:= \text{strike price of option} \\
r &:= \text{c.c. riskless rate} \\
T &:= \text{expiry of option} \\
c_t &:= \text{ value of european call at time } t \\
C_t &:= \text{ value of American call at time } t \\
p_t &:= \text{ value of european put at time } t \\
P_t &:= \text{ value of American put at time } t \\
PV_t(x) &:= xe^{-r(T-t)} \qquad (x \text{ known at time } t) \\
FV_t(x) &:= xe^{r(T-t)} \qquad (x \text{ known at time } t)
\end{aligned}$$

Throughout we adopt the following assumption:

Assumption: No dividends are paid during the life of the option.

Observe that, because of the limited liability feature of listed shares, the share price can never be negative. In the same way, option values are non–negative:

Proposition 4.3.1 $C_t, c_t, P_t, p_t \geq 0$

Proof: Observe that the terminal payoff of an option is non–negative: Both $(S_T - K)+$ and $(K - S_T)^+$ are ≥ 0.
Suppose now that some option has negative value — say $c_t < 0$. Then you can implement the folowing arbitrage strategy:

(i) At time t, "buy" the call option, for a net positive cashflow of $|c_t|$.

(ii) At time T, obtain a net positive cashflow of $(S_T - K)^+$.

You make at least $|c_t|$ using this strategy, with no initial costs, and no risk.

The same argument works for C_t, P_t and p_t.

⊣

Proposition 4.3.2 *The value of an american option is greater than its payoff if exercised, i.e.*

$$C_t \geq S_t - K \qquad P_t \geq K - S_t$$

Proof: Suppose that $C_t < S_t - K$. In that case, buy the american call and exercise it immediately for a net cashflow of $-C_t + (S_t - K) > 0$. That's an arbitrage.

A similar argument works for american put options.

⊣

The value of an option is often decomposed as follows:

$$\text{Value of Option} = \text{ Intrinsic Value} + \text{Time Value}$$

The *intrinsic value* of an option is the value obtained by exercisisng it, *if* you could exercise it. Thus for european– and american call options, the intrinsic value at time t is $(S_t - K)^+$ european– and american put options it is $(K - S_t)+$. You can't exercise european options at time t (assuming $t < T$), but if you could, then the intrinsic value equals the payoff.

Consider now an american or european call with strike $K = 100$ when the asset price is $S_0 = 90$. The intrinsic value of the option is $(S_0 - K)^+ = 0$. However, the value of the option is not zero: There is a chance that $S_T > K$ at expiry. The *time value* of an option is derived from the fact that there is a positive probability that the underlying asset will move in a favourable direction while you wait to exercise.

Time value cannot be negative for american options, since you don't *have* to wait. For deep in–the–money european put options, time value can be negative.

Observe that you should never exercise an american option if its value is greater than its payoff: For example, if $C_t > (S_t - K)^+$, then you will receive C_t if you sell the option, but only $(S_t - K)^+$ if you exercise it. (The same goes for put options). Thus an american option should be exercised only if its value equals its intrinsic value, i.e. if its time value is zero.

To summarize:

- Option value = Intrinsic value + Time value.
- Intrinsic values are easy to calculate. hence option pricing is all about calculating time values.
- Time value is always non–negative for american options.
- Time value is always positive for out-of-the-money european options. (It may be negative for deep in-the-money european puts.)

Next, it is easy to see that american options are worth more than their european counterparts:

Proposition 4.3.3 $C_t \geq c_t$, $P_t \geq p_t$

Proof: To see this intuitively, observe that anything you can do with a european option, you can also do with the corresponding american option — and more besides! To see this via an arbitrage argument, suppose $P_t < p_t$ (say). Then you can implement the folowing strategy:

Action	Cashflow at time t	cashflow at time T
Short european put	p_t	$-(K-S_T)^+$
Long american put (hold till maturity)	$-P_t$	$(K-S_T)^+$
Deposit in bank	$-(p_t-P_t)$	$(p_t-P_t)e^{r(T-t)}$
Net cashflow	0	$(p_t-P_t)e^{r(T-t)}>0$

This strategy costs nothing at time t and will have positive cashflow at time T — an arbitrage strategy.

The same argument works for call options.

$\dashv$

One fact that we will use in the sequel, but only prove later is the following:

> *It is never optimal to exercise an american call option (on a non–dividend paying asset) before maturity. Hence the value of an american call is the same as the value of the coresponding european call.*

This is not true for puts.

Proposition 4.3.4 $S_t - PV_t(K) \leq C_t, c_t \leq S_t$

Proof: We first show that $C_t, c_t \leq S_t$. One intuitive way to see this is via the following assertion: The mere *right* to buy an asset cannot be more valuable than actually having the asset.

To see this via an arbitrage argument, suppose that $S_t < C_t$.

Action	Cashflow at time t	cashflow at time T if exercised	cashflow at time T if not exercised
Short call	$+C_t$	$-(S_T-K)$	0
Buy share	$-S_t$	S_T	S_T
Deposit in bank	$-(C_t-S_t)$	$(C_t-S_t)e^{r(T-t)}$	$(C_t-S_t)e^{r(T-t)}$
Net cashflow	0	$(C_t-S_t)e^{r(T-t)}+K>0$	$(C_t-S_t)e^{r(T-t)}+S_T>0$

This strategy costs nothing at time t and will have positive cashflow at time T, whether the option is exercised or not.

Next, we show that $c_t \geq S_t - PV_t(K)$: Indeed, suppose that $c_t < S_t - PV_t(K)$. Then $FV_t(S_t - c_t) > K$. Hence we can implement the following arbitrage strategy:

Action	Cashflow at time t	cashflow at time T if exercised	cashflow at time T if not exercised
Short share	$+S_t$	$-S_T$	$-S_T$
Buy call	$-c_t$	S_T-K	0
Deposit in bank	$-(S_t-c_t)$	$FV_t(S_t-c_t)$	$FV_t(S_t-c_t)$
Net cashflow	0	$FV_t(S_t-c_t)-K>0$	$FV_t(S_t-c_t)-S_T>0$

(Observe that $FV_t(S_t - c_t) > K \geq S_T$ when the call is not exercised.)

Since $C_t \geq c_t$, we also have $C_t \geq S_t - PV_t(K)$.

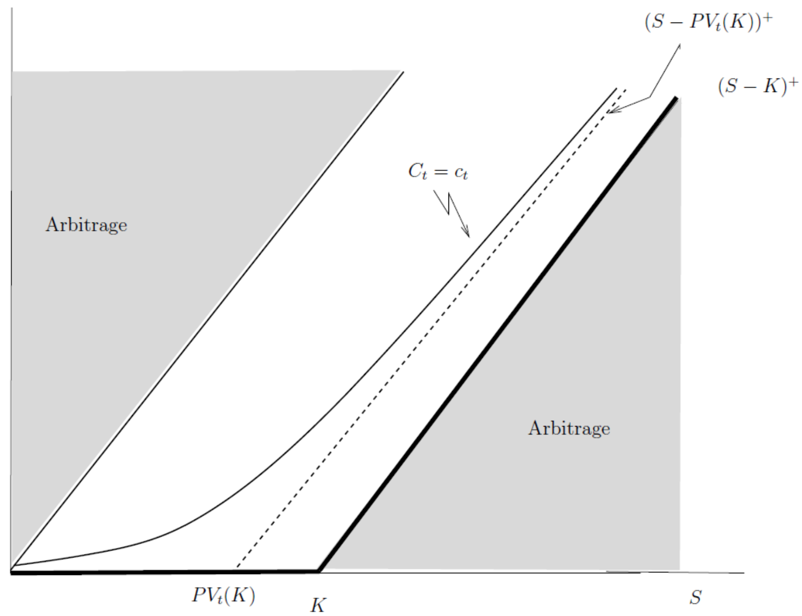

Figure 4.3: Call price as a function of stock price.

⊣

As corollary:

Corollary 4.3.5 *If $S_t = 0$, then $C_t = c_t = 0$.*

□

Proposition 4.3.6 *(a) $K - S_t \leq P_t \leq K$ and $PV_t(K) - S_t \leq p_t \leq PV_t(K)$.*

(b) If $S_t = 0$, then $P_t = K$ and $p_t = PV_t(K)$

Proof: (a) We already know that $P_t \geq S_t - K$. If $P_t > K$, then short the american put and place the proceeds in a bank account. If the holder of the put now exercises at some time T^* (where $t \leq T^* \leq T$), take K from the bank, and receive the share. You now own a share, and still have some money in the bank. This strategy cost you nothing — it is an arbitrage. Hence $P_t \leq K$.

To see that $p_t \geq PV_t(K) - S_t$, consider the following portfolio. Buy the european put and the underlying share at time t: If the put expires in-the-money, you will have $K - S_T + S_T = K$. If the put expires out–of–the–money, you will have $0 + S_T \geq K$. Either way, your portfolio will be worth $\geq K$ at time T, and thus your portfolio is worth $\geq PV_t(K)$ at time t, i.e. $p_t + S_t \geq PV_t(K)$.

Suppose now that $p_t > PV_t(K)$. Then $FV_t(p_t) > K$, where $FV_t(p_t) := p_t e^{r(T-t)}$. Now consider the following strategy:

Action	Cashflow at time t	cashflow at time T if exercised	cashflow at time T if not exercised
Short the put	p_t	$-(K - S_T)$	0
Deposit in bank account	$-p_t$	$FV_t(p_t)$	$FV_t(p_t)$
Net cashflow	0	$S_T - K + FV_t(p_t) > 0$	$FV_t(p_t) > 0$

This strategy costs nothing at time t and will have positive cashflow at time T, whether the put is exercised or not.

(b) The most that a put option can ever payoff is its strike K, since $(K - S)+ \leq K$. Thus if $S_t = 0$, then the holder of an american put should exercise immediately, in order to receive K.

For european puts, we have just shown that always $p_t \leq PV_t(K)$, i.e. that $FV_t(p_t) \leq K$. Now suppose that the inequality is strict, i.e. that $p_t < PV_t(K)$, and that $S_t = 0$. Then proceed as follows:

(i) Borrow a cash amount p_t and use the funds to buy a european put.

(ii) Also buy a share (which is free).

(iii) If the put expires in–the–money ($S_T \leq K$), receive K in return for the share. Use some of the amount K to repay the cash loan (which is possible, as $K > FV_t(p_t)$).

(iv) If the put expires out–of–the money ($S_T \geq K$), sell the share and use some of the proceeds to repay the cash loan (which is possible, as $S_T \geq K > FV_t(p_t)$).

This is an arbitrage. Hence to exclude arbitrage, it is impossible that the inequality is strict, i.e. that $p_t < PV_t(K)$ when $S_t = 0$. Since always $P_t \leq PV_t(K)$, the only possibility is that $p_t = PV_t(K)$.

$\dashv$

Finally, we expect that option prices are smooth functions of the parameters $S_t, K, r, \sigma, T-t$. Thus we have the following limit behaviour:

Proposition 4.3.7 *(a)* $c_t, C_t \to 0$ *as* $S_t \to 0$.

(b) $c_t, C_t \to S_t - PV_t(K)$ *asymptotically as* $S_t \to \infty$.

(c) $p_t \to PV_t(K)$ *and* $P_t \to K$ *as* $S_t \to 0$.

(d) $p_t, P_t \to 0$ *as* $S_t \to \infty$.

$\square$

To clarify (b) above, if S_t is much greater than K, then it is almost *certain* that the call will be exercised, i.e. that the holder will receive $S_T - K$. Thus the time t–value of such a call is nearly equal to the time t–value of a portfolio that will pay $S_T - K$ at time T. Clearly, such a portfolio has a time t–value of $S_t - PV_t(K)$.

Similarly, to clarify (d), if $S_t \to \infty$, then it is almost *certain* that the put will not be exercised, i.e. that its payoff will be 0. A portfolio that is guaranteed to have zero payoff at time T must have zero value at time t.

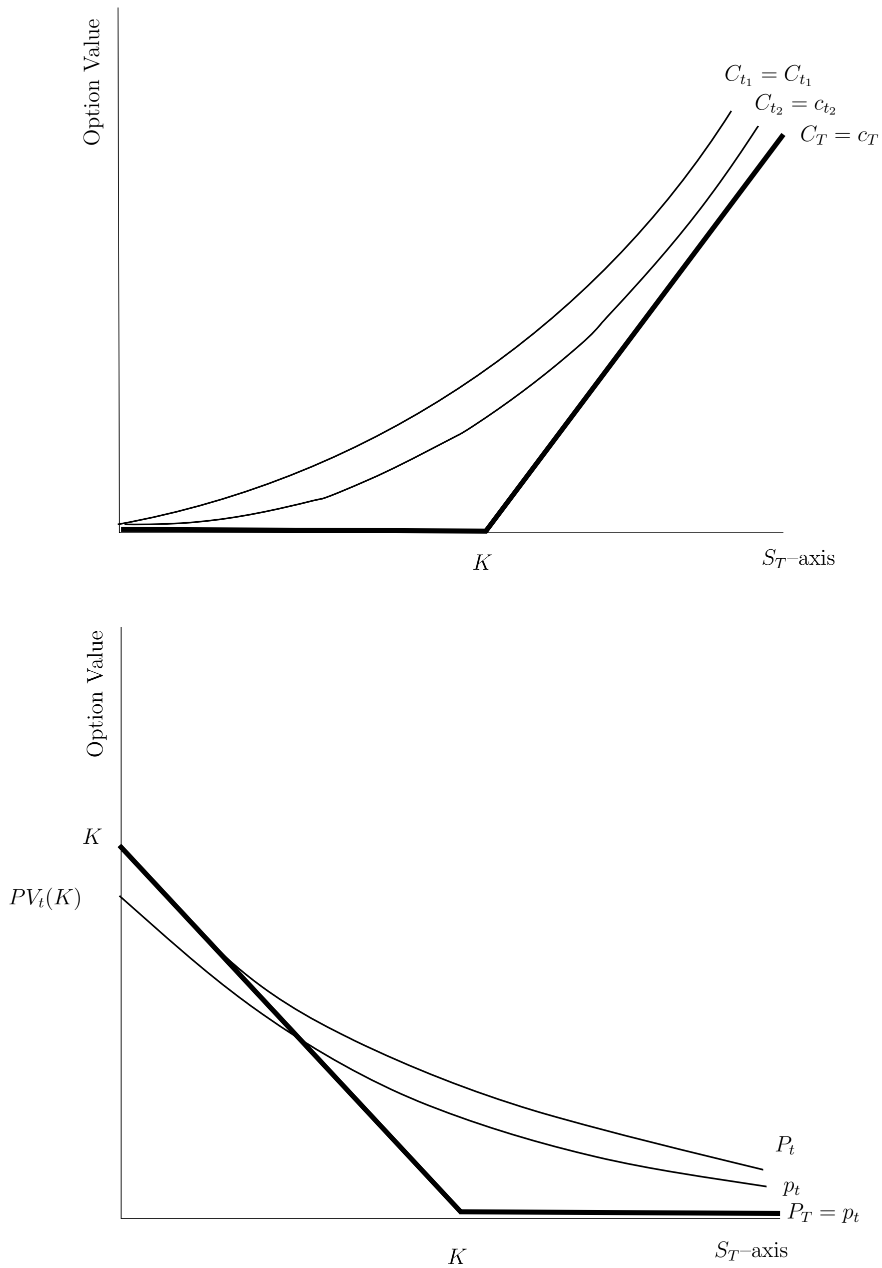

Figure 4.4: Values of american–and european options. We have $t_1 \leq t_2 \leq T$ and $t \leq T$. At expiry, the value of an option equals its payoff.

4.4 Put-Call Parity

For *european options*, assuming no interim dividends, we have the following remarkable relation:

$$\boxed{c_t + PV_t(K) = p_t + S_t}$$

In order to exercise a european call you need to have K at maturity, i.e. you need $PV_t(K)$ at time t. In order to exercise a european put, you need to have S_T at maturity, i.e. you need to have a share S_t at time t. The above put–call parity relationship says that for european options with the*same strike and maturity*,

$$\boxed{\begin{aligned} &\text{call} + \text{the cash you need to exercise the call} \\ = &\text{put} + \text{the share you need to exercise the put} \end{aligned}}$$

To see this, consider the following portfolios at time $t \leq T$

Portfolio I: Buy a european call with strike K and expiry T, and deposit $PV_t(K)$ in the bank account.

Portfolio II: Buy a european put option with strike K and expiry T, and a share S.

Portfolio I has a value of $c_t + PV_t(K)$ at time t, whereas Portfolio II has a value of $p_t + S_t$ at time t.

At expiry, there are two possibilities:

- If $S_T \geq K$, then the call is exercised, but the put is not.
  Portfolio I now has value $K + (S_t - K) = S_T$.
  Portfolio II has value $0 + S_T = S_T$.
- If $S_T \leq K$ then the put is exercised, but the call is not.
  Portfolio I now has value $0 + K = K$.
  Portfolio II now has value $K - S_T + S_T = K$

We thus see that at expiry T, Portfolios I and II are guaranteed to have the same value: If $S_T \geq K$, bot portfolios have value S_T, whereas if $S_T \leq K$ both portfolios have value K.

Since Portfolios I and II are guaranteed to have the same values at time T, they must have the same value at time t: If not, buy the cheaper portfolio and short the more expensive portfolio...

Here is another way to see this: First, observe that the payoff of a portfolio consisting of long one european call with strike K and short one european put with strike K is equal to the payoff of a long forward contract with delivery price K:

$$C_T - P_T = (S_T - K)^+ - (K - S_T)^+ = S_T - K$$

This follows because $x^+ - (-x)^+ = x$, as you can easily verify. Now recall that a long forward contract on S with delivery price K and maturity T has a value of $S_t - PV_t(K)$ at time t. Since the portfolio of options has the same value as the forward contract at time T, they must have the same value at time t, or else there will be arbitrage. thus

$$c_t - p_t = S_t - PV_t(K)$$

and the put–call parity relation follows again.

Thus if you have calls at your disposal, then you don't need puts: A long position in a put option can be replicated by a long position in a call option and cash, and a short position in the underlying share:

$$p_t = c_t + PV_t(K) - S_t$$

Exercise 4.4.1 1. Quick: For what value of K is the value of a european call equal to the value of a european put (both with strike K and maturity T)?

2. Show that if a dividend D is expected during the life of the option, then the put-call parity relation becomes

$$p_t + S_t = c_t + PV_t(K) + PV_t(D)$$

3. Show that if a stock pays a continuous dividen yield q, then the put–call parity relation becomes

$$p_t + S_t e^{-q(T-t)} = c_t + Ke{-}r(T-t)$$

4. We introduce *binary* options (also known as *digital* options.). A binary call option is a derivative security on an underlying stock S, If it has stike K and maturity T, then the option will have a payoff

$$\text{Payoff} = I_{\{S_T \geq K\}} = \begin{cases} 1 & \text{if } S_T \geq K \\ 0 & \text{else} \end{cases}$$

Similarly, a binary put option has

$$\text{Payoff} = I_{\{S_T < K\}} = \begin{cases} 0 & \text{if } S_T \geq K \\ 1 & \text{else} \end{cases}$$

Derive a put–call parity relation for binary options.

□

4.5 Early Exercise of American Options

As we have already stated, the value of an american option is always greater or equal to its intrinsic value. An american option should never be exercised if its value is strictly greater

than its intrinsic value i.e. if $C_t > (S_t - K)^+$ for a call, or $P_T > (K - S_t)^+$ for a put. The reason is simple: If $C_t > (S_t - K)+$ then it is better to *sell* the call than it is to exercise it.

We also that the value of an american option is $\geq$ that of its european counterpart: $C_t \geq c_t$ and $P_t \geq p_t$. Now, e.g., the difference $P_t - p_t$ is the additional cost that must be paid (over and above the value of the european put) to allow the holder to exercise the put early, i.e. it is an *early exercise premium*. Valuing american options is more difficult than valuing european options because it is hard to value this early exercise premium. It depends on all possible exercise strategies that the holder of the option can devise.

Before we continue, we prove a remarkable result due to Robert Merton, which shows that for american alls on non–dividend paying assets the early exercise premium is zero:

Theorem 4.5.1 *It is never optimal to exercise an amercan call on a non–dividend paying asset before maturity.*

Proof: Observe that if $t < T$, then $PV_t(K) < K$. Now

$$C_t \geq c_t \geq S_t - PV_t(K) > S_t - K$$

and thus

$$C_t > S_t - K \qquad \text{when } t < T$$

Hence the value of an american call is greater than its intrinsic value at any time before maturity.

⊣

Corollary 4.5.2 *The values of an american call and a european call (with the same strike and maturity) on a non–dividend paying asset coincide:*

$$C_t = c_t$$

Proof: An american call will be held until maturity. There is no extra value in being allowed to exercise an american call before maturity, so no rational investor would pay for this right. Hence its payoff is $(S_T - K)^+$, the same as for a european call.

⊣

Qualitatively, there are two reasons why an american call should not be exercised before maturity:

(1) **Insurance:** Options can act as insurance against the underlying asset rising or dropping in value. For example, an investor who owns a stock can protect himself against dropping prices by purchasing a put. On the other hand, an investor who is bullish on a certain stock may prefer to buy a call option, rather than the stock itself. In that way, if stock prices rise, the option will provide a positive payoff. But is stock prices fall, the investor does not suffer too great a loss: He simply discards the option.

Intuitively, long–term insurance costs more than short–term insurance. Exercising an american option early means giving up insurance. This applies to both calls and puts.

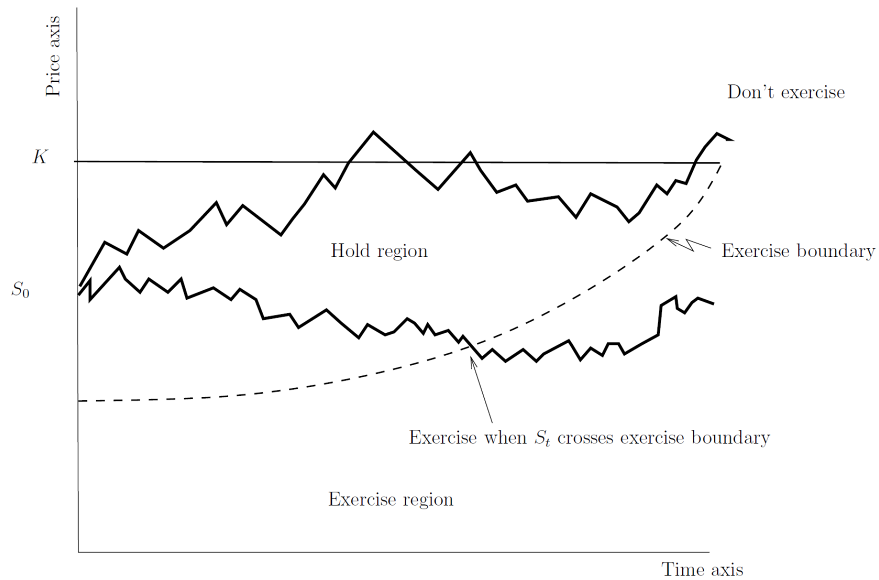

Figure 4.5: Early exercise boundary of an american put option.

(2) **Interest:** Paying K now is worse than paying K later. The holder of an american call must pay K when she exercises, and she would prefer to pay K as late as possible.

For the holder of an american put, the oppposite is true: Such an investor will *receive* K when she exercises, and she will prefer to receive K sooner, rather than later.

These two forces are *aligned* for american call options: Both insurance and interest will make the holder of such a call exercise as late as possible, i.e. at maturity. For put options, however, the forces oppose each other: It may be advantageous to forego additional insurance in order to receive K early. There is therefore a tug–of–war between insurance and interest.

For similar reasons it may be optimal to exercise an american call on a dividend paying asset: If the holder waits until maturity, he may miss out on the dividend.

Example 4.5.3 Suppose that at time $t < T$ the stock and strike prices satisfy the relationship

$$S_t < K(1 - e^{-r(T-t)})$$

If the holder of an american put exercises at time t, she will receive $K - S_t > Ke^{-r(T-t)}$. This can be invested in a bank account to receive an amount $> K$ at time T. Since an american option can never pay more than K, it is clear that exercising at time t will be better than holding the option until maturity.

□

Exercise 4.5.4 Show that if interest rates are zero, it is never optimal to exercise an american put option before expiry (unless the underlying asset price is zero).
[Hint: use put–call parity for european options.]

□

We end with a put–call parity relationship for american options:

Proposition 4.5.5 *For american options on a non–dividend paying asset, we have:*

$$S_t - K \leq C_t - P_t \leq S_t - PV_t(K)$$

Proof: We know that

$$C_t - P_t = c_t - P_t \leq c_t - p_t = S_t - PV_t(K)$$

by put–call parity for european options, and hence $C_t - P_t \leq S_t - PV_t(K)$.

To see the other inequality, we use an arbitrage argument. Suppose that T^* is a possible exercise date, where $t \leq T^* \leq T$. Consider the following strategy:

Action	Cashflow at time t	cashflow at time T^* if $S_{T^*} < K$	cashflow at time T^* if $S_{T^*} \geq K$
Short put	P_t	$-(K - S_{T^*})$	0
Buy call	$-C_t$	0	$S_{T^*} - K$
Short stock	S_t	$-S_{T^*}$	$-S_{T^*}$
Deposit K	$-K$	$Ke^{-r(T^*-t)}$	$Ke^{-r(T^*-t)}$
Net cashflow	$P_t - C_t + S_t - K$	$K(e^{r(T^*-t)} - 1) > 0$	$K(e^{r(T^*-t)} - 1) > 0$

Since the final cashflow is positive (at time T^*), the initial cashflow must be negative (i.e. the portfolio must cost money), or else there will be arbitrage. Hence $P_t - C_t + S_t - K < 0$.

$\dashv$

4.6 Option Trading Strategies

We will not demonstrate several strategies for betting on or hedging against certain events. Throughout, we assume that the underlying is a stock, and that all options are vanilla european.

The simplest trading strategies involve a single stock and/or a single option:

- *Writing a covered call:* The writer of a call option must make a payoff of $(S_T - K)^+$ at time T, and since S_T might turn out to be very large, the writer faces the risk of extremely large losses. The simplest way to hedge against such a loss is simply to own the share. Thus a covered call consists of a short position in a call option and a long position in the underlying share.

  If the writer of the call does not own the underlying stock, the call is said to be *naked*. Naked option positions may require greater initial margin on some option exchanges.

- A *protective put* is a long position in both a stock and a put option. While the value S_T of the stock at maturity might become very small, the value of the protective put portfolio can never be less than the strike K.

A *spread* is a position in two or more options of the same type, i.e. two or more calls or two or more puts.

- A *vertical spread* uses options of the same maturity, but with different strike prices.

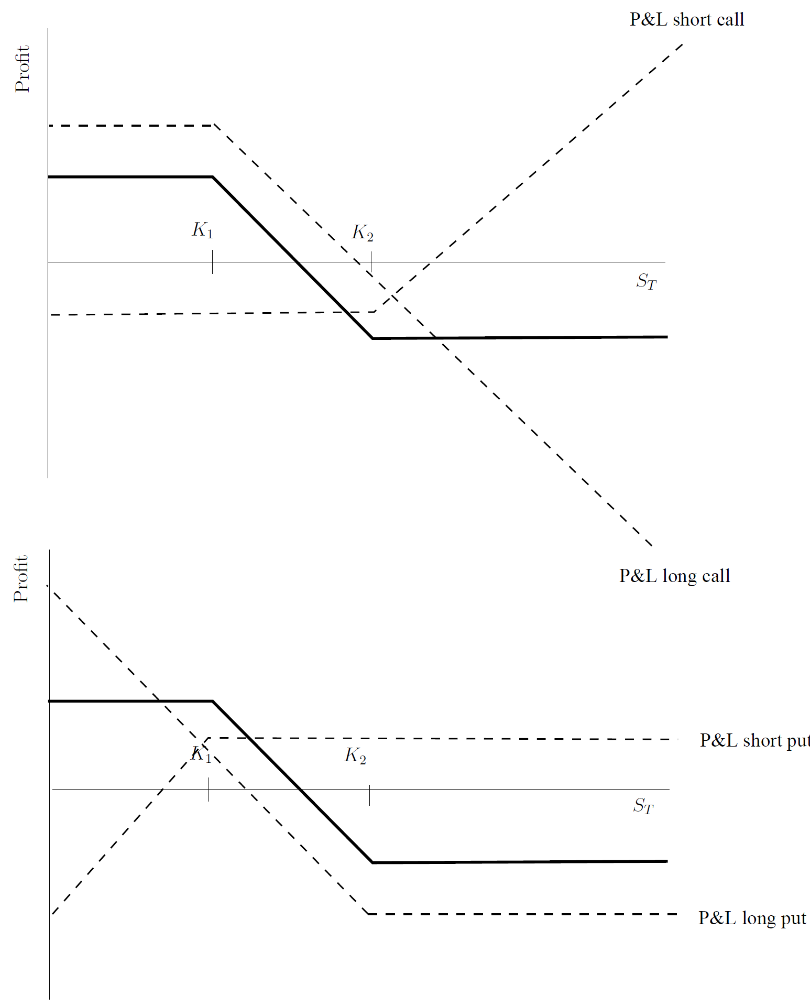

Figure 4.6: Profit diagrams for a vertical bear spread constructed from call options (top) and put options (bottom).

- A *horizontal spread* (or *calendar spread*) uses options with the same strike, but different maturities.
- A *diagonal spread* uses options with both different strikes and different maturities: It is in between horizontal and vertical.

This terminology reflects the way options are quoted in the newspapers: Strike prices are listed vertically down the page, whereas maturities are listed horizontally across the page.

Example 4.6.1 A *bear spread* is a vertical spread apprporiate for an investor who believes or fears that stock prices will decrease, while limiting losses if the belief turns out to be unfounded. In can be constructed by going long a call with strike K_1 and going short a call with strike $K_2 > K_1$. The profit diagrams are shown in Figure 4.6.

Alternatively, one can construct a bear spread with exactly the same profit or loss by going short a put with strike K_2 and long a put with strike K_2, where $K_1 < K_2$.

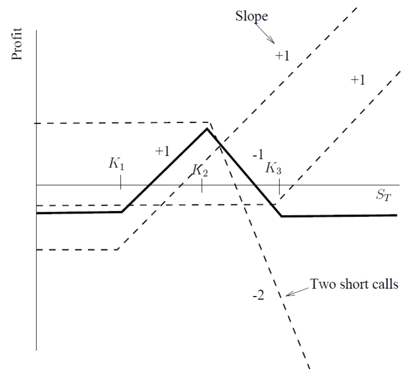

Figure 4.7: Profit diagram for a butterfly spead constructed from call options.

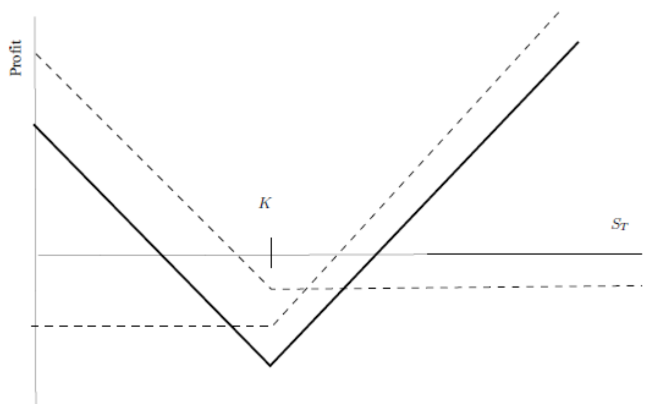

Figure 4.8: Profit diagram for a straddle.

Similarly, a vertical *bull spread* is for investors who believe or fear that stock prices will increase. It is simply a short bear spread. The profit and loss is simply the negative of the profit and loss for the corresponding bear spread.

□

Example 4.6.2 A butterfly spread is a vertical spread consiting of four options, all with the same maturity but with three strike prices, K_1, K_2, K_3 where $K2 := \frac{1}{2}(K_1 + K_3)$. A long butterfly spread consists of a long call with strike K_1 two short calls with strike K_2 and a long call with strike K_3. It is therefore equivalent to a combination of a bull spread (long call K_1, short call K_2) and a bear spread (long call K_3, short call K_2). See Figure 4.7.

Can you see how to construct a butterfly using only puts?

□

Example 4.6.3 Suppose that you know that the price of a stock is going to move at some future date, but you don't know which way. This could be because an oil company that is

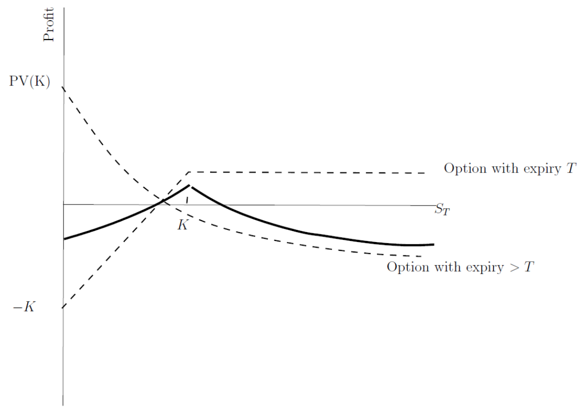

Figure 4.9: Profit diagram for a calendar spread made up of two puts, at the expiry of the short-term option.

doing expensive explorations is expected at that date, or because the results of a closely–fought election are expected at that date. At this stage you don't know if the news is going to be good (oil has been found; business–friendly party comes to power) or bad (no oil found; anti–business party comes to power). Once the news is made public, you expect that the share price will either rise dramatically, or drop dramatically.

A tool for benefitting from this is the *straddle*. Suppose the current stock price is S_0. Go long both a call and a put with strike $K \approx S_0$. If $S_T \geq K$ the call expires in-the-money, whereas if $S_T < K$, the put expires in-the-money. See Figure 4.8.

□

There are a few other combinations that are available: Draw the profit diagrems of the following strategies yourself 9or look them up), and figure out under what circumstances their use is appropriate.

- A *strangle* is a long put with strike K_1 and a long put with strike K_2, where $K_1 < K_2$.
- A *strip* is a strategy consisting of long one call and two puts, with the same strike and expiry. A *strap* is long two calls and one put, with the same strike and expiry.
- A *condor* strategy consists of long one call and one put, both equally out of the money, and short one call and one put, both equally out of the money, but further out than the options purchased. All options have the same maturity.
- A seagull consists of a short call and a short put, both equally out of the money, andan at–the–money call. All options have the same maturity.

Finally, in Figure 4.9 we show the time T–value of a calendar spread consisting of short a put with maturity T and long a put with maturity $> T$. Both options have the same strike K.

Chapter 5

Fixed Income Securities

The term “fixed income security” originally referred to instruments that pay a fixed rate of interest, usually fixed coupon bonds. A *fixed coupon bond* is a loan that pays a fixed rate of interest on fixed dates until the bond matures, when the loan is repaid. The nominal value of the loan is called the *principal* or the *face value* and the interest payments are called *coupons*.

Example 5.0.4 The R204 is a SA government bond with a face value (or principal amount) of R1 million and a 8% coupon that matures on the 21$^{\text{st}}$ of December 2018. Since coupon payments are made semiannually in South Africa, the bond holder will receive 4% of R1 million, twice a year (i.e. R40 000 on the 21$^{\text{th}}$ of June and the 21$^{\text{st}}$ of December) until 2018. On 21 December 2018, the bond holder will receive R 1 040 000 — i.e. the principal plus the final coupon.

The R157 bond pays a 13.5% semiannual coupon. It matures on the 15th of September 2014/2015/2016. On each of these three dates it will pay $\frac{1}{3}$ of the principal amount. It will continue to pay a 13.5% coupon on the remainder of the principal until 15/09/2016.

□

A bond is said to trade *at par* if its price is (approximately) equal to its face value. If its price is above (below) its face value, the bond is said to trade at a premium (discount).

Even though the cashflows of the bond are known exactly in advance, bond prices can oscillate wildly. The above definition of “fixed income securities” is no longer as narrowly

Bond Code	Maturity	Coupon	Comp. Bond	BP Spread	MTM	All IN Price	Clean Price	Accrued Interest	High Yield	Low Yield	Return	Duration	Mod. Duration	Delta	RPBP	Convexity
E170	2020/08/01	13.50	R203	80	6.5050	147.3911	141.1034	6.2877	6.6200	6.4400	0	5.1616	4.9990	-7.3681	736.8061	34.0858
R157	2015/09/15	13.50			5.3200	124.6477	120.0244	4.6233	5.4050	5.2900	0	2.2655	2.2068	-2.7507	275.0708	6.4855
R179	2013/08/01	10.00	R201	-30	4.7600	107.3668	102.7092	4.6575	4.8150	4.7250	0	.5148	.5028	-.5399	53.9876	.5090
R186	2026/12/21	10.50			7.1200	130.3453	129.5398	0.8055	7.3900	7.0850	-.02	8.3302	8.0438	-10.4847	1048.4724	90.7841
R189	2013/03/31	6.25			-.8500	217.1916	213.2313	3.9603	-.8500	-.8500	0	.1968	.1976	-.4292	42.9152	.0781
R201	2014/12/21	8.75			5.0600	107.3477	106.6764	0.6712	5.1150	5.0250	0	1.8053	1.7608	-1.8902	189.0154	4.0773
R203	2017/09/15	8.25			5.7050	113.1127	110.2874	2.8253	5.8400	5.6600	-.01	3.8986	3.7905	-4.2875	428.7520	17.9816
R204	2018/12/21	8.00			5.9500	110.7163	110.1026	0.6137	6.0850	5.9050	-.01	4.8651	4.7245	-5.2308	523.0840	27.4876
R206	2014/01/15	7.50			5.0900	102.3637	102.3020	0.0616	5.1350	5.0350	.01	.9738	.9497	-.9721	97.2121	1.3731
R223C	2013/03/15	0.00	R201	-43.5	4.6250	99.2954	99.2954	0.0000	4.6800	4.5900	0	.1559	.1523	-.1513	15.1270	.0464
R225C	2014/03/15	0.00	R201	-14	4.9200	94.5422	94.5422	0.0000	4.9750	4.8850	0	1.1547	1.1270	-1.0655	106.5465	1.8200
R227C	2015/03/15	0.00	R157	6	5.3800	89.1909	89.1909	0.0000	5.4650	5.3500	0	2.1547	2.0983	-1.8715	187.1451	5.4243
R229C	2016/03/15	0.00	R157	-18	5.1400	85.2056	85.2056	0.0000	5.2250	5.1100	0	3.1547	3.0757	-2.6206	262.0628	10.9589
SZ18	2015/09/30	12.50	R157	55	5.8700	120.0779	116.3108	3.7671	5.9550	5.8400	0	2.3262	2.2599	-2.7136	271.3648	6.7253

Figure 5.1: SA bonds/gilts on 2013/01/15. Source: Sharenet.co.za

applied, however, and now includes many debt instruments whose promised cashflows are well–defined in advance, but not necessarily known in advance, i.e. far from fixed.

5.1 Introductory Notions

In this section we introduce some basic ideas which are necessary to understand fixed income markets.

5.1.1 Interest Rates

There are many types of quoted interest rate: mortgage rates, deposit rates, prime overdraft rates, etc. These rates fluctuate with the state of the economy, with perceived risk, tax legislation, etc. For our purposes the most important are

- **Treasury Rates:** Rates at which a government can borrow in its own currency. Since governments are amongst the most prolific borrowers, the market in government debt is often quite large and quite liquid. Such government debt is, moreover, regarded as risk–free: If the government is unable to meet its debt obligations, it can always raise taxes or simply print more money. Hence these are important benchmark rates are also referred to as *risk–free rates*.

- **LIBOR Rates:** LIBOR stands for *London Interbank Offer Rate*. The *interbank market* is, as the term suggests, a money market between banks, allowing banks to borrow from each other and the central bank. This funnels excess liquidity to places were it is needed, significantly reducing liquidity risk. Thus LIBOR rates are rates at which large banks (typically with a AA–credit rating) are able to borrow from each other for short–term periods (overnight – 12 months). If 3–month LIBOR is 5.675%, then a large bank is willing to lend at an annualized rate of 5.675% over a 3–month period. Similarly, LIBID stands for *London Interbank Bid Rate*, the rate at which a bank is prepared to accept deposits (borrow). LIBOR rates are higher than Treasury rates because they are not risk–free. Nevertheless, traders often prefer to use LIBOR rates as a proxy for "actual" risk–free rates, rather than Treasury rates. This is mainly because certain regulatory policies increase the demand for Treasury instruments, thereby increasing their price, and thus decreasing the associated Treasury rates (yields).

  Government monetary policy is primarily transmitted to the economy via the interbank market. Any manipulation of central bank rates (the rate at which banks can borrow from the central bank, e.g. the Federal Reserve in the USA, or the South African Reserve Bank in SA) is quickly transmitted to the interbank market, and from there to the economy at large.

  In the Eurozone, EURIBOR (the *Euro Interbank Offered Rate*) plays an analogous role, whereas in South Africa there is JIBAR (the *Johannesburg Interbank Agreed Rate*).

- **Repo Rate:** A *repo* or *repurchasing agreement* is a contract whereby a one party agrees to sell securities it owns now, and to buy them back a short while later at a slightly higher price. (Overnight and 7-day repos are very common). In essence, this is very much like a loan with the securities as collateral. The difference between the sell price and the buy–back price implies an interest rate, the repo rate.

In South Africa, the repo rate refers to repurchase transactions between the reserve bank and its clients, the local banks. Thus it is, in effect, the rate at which local banks can borrow from the reserve bank.

5.1.2 International Debt Securities Markets

The following table should give a feel for the magnitude of the international debt markets:

Country	**Domestic**	**International**	**Total**
All	70 148	29 711	99 859
USA	26 391	6 962	33 353
Japan	14 051	183	14 234
Germany	2 802	2 164	4 966
Brazil	1 568	142	1 710
South Africa	218	35	253

Amounts outstanding in billions of US$
Source: BIS Quarterly Review, September 2012

As can be seen from the table, the USA comprises about 33% of the global debt market, and South Africa about 0.25%.

Domestic debt is further subdivided into 3 categories:

Public	*Debt issued by central government* Government bonds (sovereign bonds) Government agency bonds, guaranteed bonds State and local bonds, municipal bonds
Corporate	*Debt issued by non–financial corporations* e.g. industrial and utility companies Commercial paper, Banker's Acceptances Debentures, Certificates of deposit Corporate bonds
Financials	*Debt issued by financial corporations* e.g. investment banks, insurers

Government bonds are called *Treasuries* in the USA and *Gilts* in the UK. In the USA, one distinguishes between *Treasury bills* (which are short maturity (≤ 1 yr.) zero coupon bonds), *Treasury notes* (2–10 yr. maturity fixed coupon bonds) and *Treasury bonds* (10–30 yr. maturity fixed coupon bonds). These are auctioned to a select group of institutional investors at regular intervals, and are then traded on the secondary market. The most recently issued debt securities (called *on–the–run* or *current* securities) are the most actively traded. Although most securities are obviously *off–the–run*, these generally account for less than 25% of all trading activity.

Another way of subdividing he debt market is into the *money market* and the *bond market*.
Money market — securities with original maturity ≤ 1 year.
Bond market — securities with remaining life > 1 year.

The bond market can be subdivided as follows:
Domestic bonds — floated by a domestic borrower in domestic currency

(e.g. Eskom in RSA in Rands)
Foreign bonds — floated by a foreign issuer in domestic currency
(e.g. Old Mutual in London in £.)
Eurobonds — floated by a foreign or resident issuer in foreign currency
(e.g. Old Mutual or IBM in Johannesburg in \$.)
Foreign bonds and Eurobonds together comprise the *international bond market*. Foreign bonds often have colourful and exotic names, e.g. Yankee bonds (USA), Samurai (Japan), Bulldog (UK).

The Eurobond market developed in the 1960's as a response to tax controls in the USA (originally referring to dollar dominated debt in Europe). Even after these controls were abolished, the Eurobond market kept growing because it conferred the following advantages:

- Payments in Eurobonds are not taxed (in the USA). Investors are therefore willing to accept a lower coupon rate, i.e. a corporation can raise funds at lower rates by issuing Eurobonds.
- Eurobonds do not need to be registered with the SEC[1], an uncommonly slow process. Hence the Eurobond market offers a fast and efficient environment in which to raise capital.
- Eurobonds are bearer instruments, i.e. owners do not need to register. the confidentiality of owners is therefore protected.
- Eurobonds have lower credit risk than US corporate bonds.

Of course, some bonds cannot be fitted comfortably into any of the above categories, but exhibit a hybrid structure: Dual currency bonds, for example, pay the principal in one currency, but the coupon in another.

5.1.3 Types of Fixed Income Securities

We give a brief description of some commonly used instruments.

- A **money market account** is simply a bank account which offers the prevailing (and thus constantly changing) rate of interest.
- **Annuities** pay a constant amount at regular intervals until maturity. These constant payments thus include both the interest and part of the principal. The gradual repayment of the principal is called *amortization*.
- **Fixed coupon bonds** have already been defined. They pay a fixed percentage (the coupon) of the principal at regular intervals and the principal at maturity. Treasury notes and bonds are examples of this.
- **Zero coupon bonds**, also known as *discount bonds*, do not pay any coupons, but only the principal at maturity. These are a form of short term debt. Some examples are Treasury bills (which are issued with maturities of 13, 26 and 52 weeks), commercial paper (unsecured short–term debt issued by corporations), banker's acceptances (a letter of credit issued by a bank on behalf of a corporation, i.e. a form of secured debt).

[1]Securities Exchange Commission, a US regulatory agency.

Negotiable Certificates of Deposit (NCD's) are another example. A certificate of deposit is issued in connection with a fixed deposit made at a bank. Substantial penalties are imposed on early redemption of a fixed deposit, and the resulting reluctance of investors to commit funds for a fixed period led to the invention of NCD's. These are bearer instruments that can be traded in the secondary market atmarket prices. Thus an investor who has committed funds to a fixed–term deposit at a bank can nevertheless exit this commitment, without exacting the penalties, by selling the NCD. The new owner of the NCD will now get the original deposit (plus the interest) from the bank at maturity.

As stated, zero coupon bonds typically have short matrurities, i.e. one year or less. However, banks sometimes create *synthetic* zero coupon bonds with much greater maturities. This is done stripping the coupon payments of a coupon–bearing bond from the principal, and selling these separately. Thus

$$\text{Coupon Bond} = \text{Zero Coupon Bond} + \text{Annuity}$$

For example, the US Treasury inaugurated its STRIPS programme (where "STRIPS" stands for *Separate Trading of Registered Interest and Principal Securities*) in 1985.

- **Consols** or **perpetual bonds** are fixed coupon bonds which mature at ∞, i.e. they pay a fixed percentage of the principal at regular intervals for all time, but the principal is never repaid.

- **Floating rate notes** (FRN's) are bonds that pay a variable coupon at regular intervals. This variable coupon is generally linked to some market–observable reference rate. For example, a 2–year \$1 million FRN paying LIBOR +2% semiannually in arrears, will at 6–month intervals pay

$$\frac{1}{2} \times 1\,000\,000 \times [\text{LIBOR} + 2\%]$$

where LIBOR is the variable 6–month LIBOR rate observed in the market 6 months before the coupon payment. Suppose, for example, that 6–month LIBOR evolves as in the 2nd row of the table below. Then cashflows associated with the FRN are shown in the 3rd row:

Time (months)	$t = 0$	$t = 6$	$t = 12$	$t = 18$	$t = 24$
6–month LIBOR	10%	11%	9%	10%	8%
Payments	-\$1 000 000	\$60 000	\$65 000	\$55 000	\$1 060 000

The dates $t = 0, 6, \ldots$ etc. are called the *reset dates*.

- **Structured notes** are a class of debt instruments with more complex pay–off patterns, possibly tailored to an investor's requirements. As an example, we mention *inverse floaters*. These have coupon payments that vary inversely with a reference rate. For example, the coupon might be $(15\% - \text{LIBOR})^+$ or $(20\% - 2\,\text{LIBOR})^+$. Inverse floaters can be used to hedge against falling interest rates.

- **Inflation–linked bonds** (also known as ILB's or *linkers*) are bonds where the coupon payments are linked to the prevailing inflation rate. For example, coupon payments may be 2% + inflation rate.

The R210 is a South African ILB that matures on 31/03/2028. It pays a coupon of 2.6% on the *capital value*, where the capital value is defined to be Principal Amount × Consumer Price Index.

Many of the above types of debt security can also be issued with built–in optionality.

Callable bonds	—	can be called back by the issuer at fixed prices on fixed dates. For example, ABCorp. floats a 10–year 15%–fixed coupon bond. After 5 years, its cost of capital is in the region of 6%, but ABCorp. is still paying 15%. If the bond has a call feature built in, ABCorp. can call back the bond, and float a new issue with a 6% coupon instead.
Puttable bonds	—	can be put back to the issuer at fixed prices on fixed dates. For example, investor X buys a 20–year 10%–fixed coupon bond from AA-rated ABCorp. After 10 years, this credit rating has migrated to B-, and the market requires a credit spread of 4%. The bond is now worth very little, but if it has a put feature built in, then investor X can sell the bond back to ABCorp. at a reasonable price.
Convertible bonds	—	can be converted to equity at a predetermined *conversion ratio* on predetermined dates. For example, a bond with a conversion ratio of 3 allows the bond to be converted to 3 shares.

A long position in a callable bond is obviously equivalent to a long fixed–coupon bond (with the same coupon and maturity) and a short call on this fixed coupon bond. Hence a callable bond is cheaper than its non–callable counterpart. Similarly a long position in a puttable bond is equivalent to a position which is long both a fixed–coupon bond and a put option. Hence a puttable bond trades at a premium to its non–puttable counterpart. A convertible bond is nothing more than an ordinary bond plus a warrant[2].

Other kinds of optionality are possible as well. For example, FRN's often have a maximum coupon rate. Such a capped FRN is equivalent to a standard FRN and a short cap[3]. Mortgage–backed securities are loans collateralized by property, and there is always a risk that home owners will repay their loans early. In effect, this means that home owners are long an american call option on the debt.

5.2 Pricing of Fixed Income Securities

Fixed income securities are valued by discounting their cashflows at the appropriate discount rate.

5.2.1 Bond Pricing Basics

We first consider the pricing of a zero coupon bond. Let $B(t,T)$ be the time–t value of a zero coupon bond with principal 1 and maturity T . With this price, we can associate several rates:

[2]For the moment, a warrant is a kind of call option on a share.

[3]A cap is a call option on an interest rate.

- **Discount rate i_d:** This is defined by

$$B(t,T) = 1 - \frac{i_d(T-t)}{360}$$

where T is measured in *days*. The number 360 arises from a *day count convention* that a year has 360 days. Different markets have different day-count conventions.

- **Simple rate i_s:** This is defined by

$$B(t,T) = \left(1 + \frac{i_s(T-t)}{365}\right)^{-1}$$

where T is in days. Here a 365-day year day count convention is seen to hold.

- **Discretely compounded rate $i_{d,n}$:** This is defined by

$$B(t,T) = \left(1 + \frac{i_{d,n}}{n}\right)^{-n(T-t)}$$

where T is in years.

- **Continuously compounded rate i_c:** This is defined by

$$B(t,T) = e^{-i_c(T-t)}$$

where T is in years.

Note that as the compounding frequency increases the associated rate drops, i.e.

$$i_s \geq i_{d,2} \geq i_{d,3} \geq \cdots \geq i_c$$

Intuitively, money is working harder as the compounding frequency increases, so you need a smaller rate to get the same result. Of course, your intuition (if you have any) will frequently mislead you, so you should prove that rates behave in this way. Start by showing that $f(n) = (1 + \frac{x}{n})^n$ is an increasing function of n.

For pricing purposes, the notion of a **zero rate** or **spot rate** is invaluable. The n–year zero rate is defined to be the rate implied by the price of a (possibly synthetic) zero coupon bond that matures in n years. As stated before, most zero coupon bonds have very short maturities, and hence most market–observable rates are not zero rates. Nevertheless, the spot rate curve can be (imperfectly) extracted from market data by a technique known as *bootstrapping the yield curve*.

Once the spot rate curve is known, one can use it to value all sorts of securities. For example, the (theoretical) price of a fixed coupon bond is equal to the present value of all its cashflows, discounted at the appropriate rate. Hence a fixed coupon bond is equivalent to a portfolio of zero coupon bonds.

Example 5.2.1 Suppose that we are given the following *term structure* of zero coupon bonds, i.e. the prices of zero coupon bonds with face value 1 and maturities T:

T	$\mathbf{B(0,T)}$
0.5	0.96
1.0	0.91
1.5	0.85
2.0	0.78

It is easy to see, after some simple computations, that the continuously and semiannually compounded rates implied by these prices are

T	$\mathbf{i_c}$	$\mathbf{i_{d,2}}$
0.5	8.16%	8.33%
1.0	9.43%	9.66%
1.5	10.83%	11.13%
2.0	12.42%	12.82%

Suppose that we want to price a 2–year bond with an 8% semiannual coupon and face value \$1 million. This bond will pay \$40 000 at each of $t = 0.5, 1.0, 1.5$ and \$1 040 000 at $t = 2.0$. It can therefore be seen as a portfolio of 40 000 of each of $B(0, 0.5), B(0, 1.0), B(0, 1.5)$ and 1 040 000 of $B(0, 2.0)$. Its price is therefore

$$P = 40\ 000[0.96 + 0.91 + 0.85] + 1\ 040\ 000[0.78] = 920\ 000$$

On the other hand, we can also calculate the price as

$$P = 40\ 000[e^{-0.0816\times 0.5} + e^{-0.0943\times 1.0} + e^{-0.1083\times 1.5}] + 1\ 040\ 000e^{-0.1242\times 2.0}$$

or as

$$40\ 000\left[\left(1+\frac{0.0833}{2}\right)^{-1} + \left(1+\frac{0.0966}{2}\right)^{-2} + \left(1+\frac{0.1116}{2}\right)^{-3}\right] + 1\ 040\ 000\left(1+\frac{0.1282}{2}\right)^{-4}$$

□

Price is not the best way to compare bonds, however. Suppose that bond A is a 2–year bond paying a semiannual coupon of 8% which costs \$920 000. Suppose further that bond B is a 3–year bond paying a semi–annual coupon of 10% which costs \$930 000. Which bond is best, if both have the same credit quality? The price alone tells you nothing. What you need to know is the rate of return on the capital you invest. The bond with the highest rate of return is the better one (other things being assumed equal). The internal rate of return of a bond is called its *yield*.

The **yield** (or **yield–to–maturity**, YTM) of a bond y is defined to be that number y which gives the correct price of the bond if its cashflows are discounted at a rate of y.

Example 5.2.2 To find the semiannually compounded yield of bond A above, we have to solve the following equation for y:

$$920\ 000 = \sum_{k=1}^{4} \frac{40\ 000}{\left(1+\frac{y}{2}\right)^k} + \frac{1\ 000\ 000}{\left(1+\frac{y}{2}\right)^4}$$

In this case, we have to find the roots of a fourth degree polynomial. For bond B we would have to find the roots of a 6$^{\text{th}}$ degree polynomial. In general, we cannot find the yield of an

arbitrary bond exactly, but will have to use a numerical method, such as Newton's method. Excel's *Solver* is also very handy for finding yields. Solver tells us that Bond A has a yield of 12.65%, whereas bond B has a yield of 12.89%. Thus Bond B offers the higher rate of return, and is therefore a better deal (other things being equal).

□

In the USA bonds are quoted in price, whereas in SA they are quoted in yield.

Note that if the yield of a bond increases, then its price decreases, and vice versa, i.e. price an yield are inversely related. Bonds that are perceived to carry similar risks should have approximately the same yield, because investors will prefer bonds with a higher yield, driving up their prices, and thus lowering their yields.

Note:

If the coupon rate equals the yield, then the bond trades at par.

(assuming, of course, that we are pricing at $t = 0$ or on a coupon date, just after the coupon has been paid.) This is easy to understand intuitively: The yield on a bond is essentially the rate of return demanded by the market, whereas the coupon is the rate of return offered by the bond (as a percentage of the par value). If the coupon rate equals the yield, then the rate demanded is equal to the rate offered, and the price of the bond does not need to change. If the rate demanded by the market is higher than the rate offered, then the bond is not attractive at its current price, and its price will drop, so that the coupon will form a greater percentage of the bond's new value. The opposite is true if the yield is lower than the coupon, and the bond will trade at a premium.

To see this mathematically, consider a bond with face value F and maturity T. Suppose that the coupon rate is c, paid n times annually, and that the n–times annually compounded yield is y. Then the bond price is

$$\begin{aligned} P &= F\left[\sum_{k=1}^{n} T\frac{c/n}{(1+y/n)^k} + \frac{1}{(1+y/n)^{nT}}\right] \\ &= F\left[\frac{c}{y}\left(1 - \frac{1}{(1+y/n)^{nT}}\right) + \frac{1}{(1+y/n)^{nT}}\right] \end{aligned}$$

Clearly $P \geq F$ if and only if $c \geq y$. While we are here, can you see that the price of a consol is $\frac{c}{y}F$, regardless of compounding frequency?

Apart from yield–to–maturity, there are several other useful notions of yield:

- **par yield:** (or *par bond yield*) This is the *coupon rate* that causes the bond to equal its face value. For example, to find the par yield of a bond with maturity T and a semiannually compouinded coupon, you have to solve the following equation for c:

  $$\sum_{k=1}^{2T} \frac{c}{2} B\left(0, k/2\right) + B(0,T) = 1$$

  where $B(0,T)$ is the price of a zero coupon bond with face value 1 and maturity T. The par yield will be very important later when we calculate *swap rates*.

- **current yield:** This is simply the rate of interest that a bond is currently paying, regardless of its face value or its maturity. It is defined by

$$\text{Current yield} = \frac{\text{annual coupon}}{\text{bond price}}$$

  Thus if a par = 100 15–year bond with a 7% coupon trades at 80, its current yield is $\frac{7}{80}$.

- **yield–to–call:** This applies only to callable bonds, and is the yield if the bond is called at the first call date. Suppose that a par $= F$ bond matures at time T and pays an annual coupon of c. If the bond can be called at time T^* for an amount F^* (the strike), then the yield–to–call y^* is the solution of the equation

$$P = \sum_{k=1}^{T^*} \frac{cF}{(1+y^*)^k} + \frac{F^*}{(1+y^*)^{T^*}}$$

A consol never matures, nor does it ever repay its face value. Thus you should not be surprised to find that for a consol the notions of yield–to–maturity, par yield and current yield all coincide. Prove that this is so.

5.2.2 Bond Price Quotation

The bond pricing formula for a bond that pays its coupon n times p.a.

$$P = F\left[\sum_{k=1}^{n} T\frac{c/n}{(1+y/n)^k} + \frac{1}{(1+y/n)^{nT}}\right] \qquad (*)$$

is true only if "today" is a coupon date. If this is not the case, i.e. if the date t on which the bond is sold falls strictly in between two coupon dates t_{k-1}, t_k, then the new owner of the bond will get the entire coupon c/n at t_k , i.e. the entire interest for the period $[t_{k-1}, t_k]$ is paid to the new owner. However, the new owner is really only entitled to interest for the period $[t, t_k]$; the interest for the period $[t_{k-1}, t]$ surely belongs to the original owner. This interest is incorporated into the bond price, so that

$$P_{\text{all in}} = P_{\text{clean}} + \text{AI}$$

i.e. the actual, *all–in*, or *gross*, price of the bond equals the *clean price* plus the *accrued interest*. The clean price is simply the price on the previous coupon date, and is found using $(*)$. The accrued interest is calculated on a linear basis, i.e. if an investor buys the bond on a date t between two coupon dates t_k, t_{k-1}, then the original owner is entitled to a fraction

$$\text{AI} = \frac{c}{n}(t - t_{k-1})$$

of the interest c/n for that period. This is added to what the bond would have cost on the previous coupon date (if market perceprion had been what it is today).

Example 5.2.3 US Treasury bonds are quoted in 32^{nds} of dollars, i.e. a price of 92:25 means $\$92\frac{25}{32}$. Only *clean* prices are quoted. The reason for this is that bond prices change because of two completely different reasons. The first reason for changing bond prices is changes in market perception, due to new information, economic factors, etc. However, even in absence of these, the price of a bond would change over time because of interest: A 1–year zero costing 0.90 today will cost 1.00 in a year's time, even if the market is completely static. Quoting only clean prices makes it easier to see how bond prices are changing because of market– and economic factors, i.e. the price change due to interest is removed. This makes it possible to compare today's bond price with yesterday's bond price.

Suppose now that on 7/06/2002 the $13\frac{3}{4}$ August 2004 bond is quoted at 121:13. This means that a bond with a face value of \$100 and a semiannual coupon of 13.75%, maturing on the 31^{st} of August 2004, has a clean price of $\$121\frac{13}{32}$. The coupon dates are on the 28th of February and the 31 of August. Since 99 days have passed since the last coupon date, the gross price of the bond is $121\frac{13}{32} + \frac{13.75}{2} \times \frac{99}{365} = \$123.27$.

□

Example 5.2.4 In South Africa, bonds are quoted in *yield* rather than price. This is even better than the US custom of quoting clean prices, because it allows the performance of different bonds to be compared. It makes the calculation of the all in price of a bond slightly more difficult, though hardly at all if you have a good spreadsheet.

Suppose that an institutional investor buys an R153 bond on 16/10/2002. The R153 pays a 13% semiannula coupon, has coupon dates 28/02 and 31/08, has a face value of R1 million, and matures on 31/08/2010. the quoted yield is 14.5%.

The clean price is the price on the previous coupon date, i.e.

$$P_{\text{clean}} = 1\ 000\ 000 \left[\sum_{k=1}^{16} \frac{0.065}{(1+\frac{0.145}{2})^k} + \frac{1}{(1+\frac{0.145}{2})^{16}} \right] = \text{R}930\ 308.93$$

because the there are 16 coupon dates left since the last one. The last coupon date was 46 days ago. the original owner of the bond is therefore entitled to

$$\text{AI} = 1\ 000\ 000 \left[0.065 \times \frac{46}{365} \right] = \text{R}8\ 191.78$$

The all in price of the bond is therefore R938 500.71.

□

In reality the calculation of the all in price in the above example is slightly more complicated, because the contract date (i.e. the date on which it is agreed to buy/sell a bond) is not the same as the settlement date (i.e. the date on which payment is made/received). Moreover, in SA the treasury keeps a register of bond owners, and it closes this a month prior to the payment of the coupon. If the bond is sold between the books–closed–date and the coupon date, the coupon is nevertheless paid to the original owner, not the new owner. This too, must be incorporated into the bond price.

Example 5.2.5 The South African bond pricing formula for the unrounded all–in price of a bond with nominal value R100 (and more than 6 months to maturity) is given by:

$$\text{AIP} = 100 D^{d_s/d_l} \left[c\left(x + \frac{D(1-D^N)}{1-D} \right) + D^N \right]$$

□

where:

- $D = \frac{1}{1+y/2}$, where y is the quoted yield (e.g. $y = 0.0512$ when the yield is 5.12%).
- c is the coupon rate (e.g $c = 0.04$ for a bond paying a 4% coupon).
- x is the cumex flag. Thus $x = 1$ if the next coupon is included (i.e. the settlement date for the saleof the bond is before the books closed date); else $x = 0$.
- N is the number of 6 month periods from the next coupon date to the maturity of the bond.
- d_l is the number of days from the last coupon date until the next.
- d_s is the number of days from the settlement date to the next coupon date.

□

5.2.3 Pricing FRN's and Inverse Floaters

Recall that a floating rate note is a debt contract that pays coupon linked to some market reference rate, e.g. LIBOR. The coupon rate is reset at regular intervals. For example, consider corporation ABCor which has issued a par \$1 million FRN paying quarterly coupons of (3 month) LIBOR + 1%. The 1% must be regarded as a *risk premium* that must be paid to investors in order to get them to buy the note. Big banks can borrow amongst each other at LIBOR rates, but ABCor is less creditworthy than a big bank. Hence its debt is less attractive than that of a big bank, all things being equal. In order to get the funds it needs, ABCor must ensure that all things are not equal: It must offer a higher rate of return as an inducement to take on the extra credit risk.

Suppose that ABCor has issued its FRN today, at a cost of \$1 million. Today 3 month LIBOR is 12%. Investors therefore require a return of 13% in order to invest in ABCor debt, i.e. they require an amount of \$32 500 on \$1 million over 3 months. After 3 months, the FRN pays a coupon of \$32 500, i.e. the investors are paid the interest that they require. The value of the FRN after 3 months is therefore again \$1 million (just after it has paid the coupon). Now market conditions have changed, and LIBOR is 11%. The coupon is now reset to 12%, because investors now require a return of 12% on ABCor debt (provided its credit rating has not changed). After 3 months, the FRN pays the required interest, so after 6 months, the value of the FRN is again equal to \$1 million.

Thus an FRN is very must like a deposit in a bank account at floating rates, where the interest is withdrawn from the bank account at regular intervals. Of course, if all interest paid is withdrawn, the value of the deposit equals its initial value. Hence:

On a reset date, a floating rate note trades at par[4].

Example 5.2.6 Another way to see this is as follows: Suppose an FRN (with face value 1) pays LIBOR in arrears at times $T_i = i\Delta T$ for $i = 1, \ldots, N$. Let L_i be the LIBOR rate for the period $[T_{i-1}, T_i]$. L_i is *known* at time T_{i-1}, but not before.

[4]provided the credit rating or the market's perception of the issuing corporation's relative risk is unchanged.

At time T_N the FRN matures and pays $1 + L_N\Delta T$. Thus its value at time T_{N-1} is $\frac{1+L_N\Delta T}{1+\Delta L_N\Delta T} = 1$ — we can discount to time T_{N-1} because we know the prevailing rate L_N at time T_{N-1}. * Thus at time T_{N-1} the value of the FRN, after paying the coupon $L_{N-1}\Delta T$, is 1. It follows that the value of the FRN at time T_{N-1} just before paying the coupon is $1 + L_{N-1}\Delta T$. Discounting, we see that its value at T_{N-2} is $\frac{1+L_{N-1}\Delta T}{1+L_{N-1}\Delta T} = 1$ — we can discount to time T_{N-2} because we know the prevailing rate L_{N-1} at time T_{N-2}.

Continuing backwards inductively in this way, we see that the value of the FRN is 1 on each coupon date (just after the coupon has been paid).

□

If this is kept in mind, valuing a floating rate not is a simple affair. Simply discount the par value plus the next coupon from the next coupon date to today. It follows that a FRN trades close to its par value

Example 5.2.7 ABCor has issued an FRN with face value \$1 million paying semiannual coupon of 6–month LIBOR in arrears. Three months ago, 6–month LIBOR was 10%. The next reset date is in three months' time, at which point the value of the FRN will be equal to \$1 million. A coupon of \$50 000 will also be paid. Hence the value of the FRN today is

$$\text{FRN} = B(0, \frac{1}{4})\,[\$1\ 050\ 000]$$

where $B(0, \frac{1}{4})$ is the price of a zero coupon bond with face value \$1.00, maturity 3 months and the same credit risk as ABCor.

□

Consider now an inverse floater, paying a semiannual coupon of 18% − LIBOR on a face value of \$1 million. To value this debt security, we construct the following portfolio Π: Π consists of one inverse floater, and one par \$1 million FRN paying a LIBOR semiannually. It is easy to see that Π is equivalent to a fixed coupon bond with a face value of \$2 million paying a 9% semiannual coupon. Thus

$$\begin{aligned} & \text{IF with face } F \text{ and coupon LIBOR} - x\% \\ = \; & \text{Fixed coupon bond with face } 2F \text{ and coupon } \frac{x}{2}\% \\ & - \; \text{FRN with face } F \text{ paying LIBOR} \end{aligned}$$

Similarly, an inverse floater paying 20%-2 LIBOR can be regarded as a portfolio long a fixed coupon bond paying a coupon of $\frac{20}{3}$% and short two FRN's paying LIBOR.

5.3 The Term Structure of Interest Rates

The interest rate which you must pay if you borrow for a period of one year may not be the same as the rate that you must pay if you borrow for two years. The phrase "*term structure of interest rates*" refers to the collection of (annualized) interest rates as a function of borrowing period. A graph of the interest rates versus borrowing period is called the *yield curve.* The phrase *term structure of zero coupon bonds* refers to the collection of prices of zero coupon bonds as a function of maturity.

There are many kinds of interest rates, and we have already singled out default–free zero rates as being of particular importance for the pricing of financial instruments. However, zero rates are only observable for very short maturities. In order to contsruct the yield curve of spot rates one frequently needs to to look at the par yield curve and the term structure of *forward rates*.

Example 5.3.1 Suppose that we are given the term structure of zero coupon bonds as shown in the second column of this table. Then you will easily verify that the term structure of annually compounded zero rates is given by the third column.

$\mathbf{T}$	$\mathbf{B(0,T)}$	$\mathbf{y}$
1	0.9494	5.3297%
2	0.8982	5.5148%
3	0.8470	5.6912%
4	0.7967	5.8464%

We can now calculate the par yield curve. Recall that the par yield is the coupon rate which sets the price of a coupon bearing bond equal to par. Thus to calculate c_1 the par yield of a one–year bond, we must slove

$$(c_1 + 1)B(0, 1) = 1$$

so that $c_1 = 5.3297\%$. Similarly, to find the par yield c_2 of a two–year bond, we solve

$$c_2B(0, 1) + (c_2 + 1)B(0, 2) = 1$$

which gives $c_2 = 5.5099\%$ Similary, $c_3 = 5.6780\%$ and $c_4 = 5.8230\%$. Here is a graph of the spot yield and par yield curves: Note that each "curve" consists of just 4 points. The points

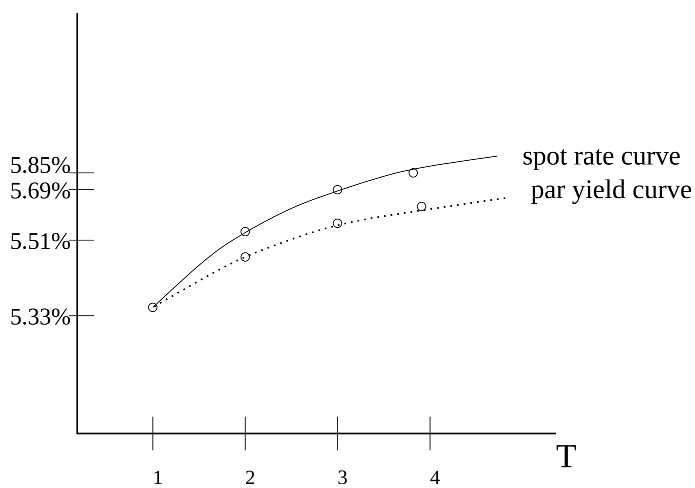

Figure 5.2: par yield and zero curves

in between are determined by some form of interpolation (in this case, using cubic splines produced by a drawing package).

□

In the example above, the term structure of zero rates is increasing. This is often, but not always the case. The yield curve may be decreasing or humped as well.

Note that if the spot curve is upward sloping, then the par yield curve lies below it, as in figure 5.3.1. The intuition behind this fact is simple: Suppose you have an n–year loan. The greater n, the higher the required interest rate, because the yield curve is increasing. If you make interim interest payments however (as a fixed coupon bond does), then those interim payments should correspond to rates that are less than the n–year rate. The par yield is a kind of average of all the required interim interest payments, and that "average" will be less than the n–year rate. Of course, exactly the reverse is true if the spot curve is downward sloping: In that case the par yield curve lies above it.

As always, intuition is important, but no substitute for mathematics: Let y_n^* denote the n–year par yield, and y_n the n–year spot rate (assuming annual coupons, and thus annual compounding). Then, by definition of $y{*}_n$,

$$\sum_{k=1}^{n} \frac{y_n^*}{(1+y_k)^k} + \frac{1}{(1+y_n)^n} = 1$$

because y_n^* is that coupon which makes an n–year bond trade at par. Now if the spot curve y_k is increasing, then

$$\sum_{k=1}^{n} \frac{y_n^*}{(1+y_k)^k} + \frac{1}{(1+y_n)^n} \geq \sum_{k=1}^{n} \frac{y_n^*}{(1+y_n)^k} + \frac{1}{(1+y_n)^n}$$

The right–hand side is the price of an n–year bond with coupon y_n^* and yield y_n, and this will trade at par when $y_n^* = y_n$. Clearly the lefthand side will be ≥ 1 if $y_n^* \geq y_n$. Since we must have LHS $= 1$, it follows that $y_n^* \leq y_n$.
If the yield curve is downward sloping, then

$$\sum_{k=1}^{n} \frac{y_n^*}{(1+y_k)^k} + \frac{1}{(1+y_n)^n} \leq \sum_{k=1}^{n} \frac{y_n^*}{(1+y_n)^k} + \frac{1}{(1+y_n)^n}$$

and the opposite is true.

5.3.1 Forward Rates

A *forward rate* is an interest rate for a future period that can be locked into today. Suppose, for example, that a corporation knows today that it will need to borrow a substantial amount in one year's time, and intends to repay this loan over two years. The corporation faces the risk that interest rates will increase over the next year. Today the 2–year rate might be 5%, but in one year's time at might be 9%, making then cost of borrowing much higher. However, the corporation can hedge by entering into a *forward rate agreement* (FRA) with a bank. This is an OTC instrument that guarantees that a certain interest rate will apply to a certain principal over a prespecified period in the future. For example, the corporation might enter into a 2–year FRA, starting in one year's time, at a forward rate of 5.5%. If interest rates do go up to 9%, this is the bank's problem: The corporation only pays 5.5%. Of course, if interest rates drop to 2%, then the bank gains, as the corporation will still have to pay 5.5%.

How are forward rates determined? They are implied by today's spot curve, using an arbitrage argument. Suppose that at time t a corporation wants to lock into an interest rate

for the period $[T_1, T_2]$, where $t \leq T_1 \leq T_2$. As usual, let $B(t,T)$ be the time–t price of a zero coupon bond with face value 1 and maturity T. We shall call such a bond a T–bond.

Now consider the following trading strategy:

Time	**Action**	**Cashflow**
t	Buy one T_1–bond Short $\frac{B(t,T_1)}{B(t,T_2)}$ T_2–bonds	$-B(t,T_1)$ $\frac{B(t,T_1)}{B(t,T_2)}B(t,T_2)$ Net cashflow at time $t=0$
T_1	Honour T_1–bond	$-\$1.00$
T_2	Redeem T_2–bonds	$\frac{B(t,T_1)}{B(t,T_2)}$

Thus this strategy requires no net outlay of capital at time t. At time T_1, you have to pay \$1.00, and you will receive \$ $\frac{B(t,T_1)}{B(t,T_2)}$ at time T_2. Thus it is as though you put \$1.00 in the bank at t_1 to receive $\frac{B(t,T_1)}{B(t,T_2)}$ at time T_2. This implies a forward simple rate L:

$$1 + L(T_2 - T_1) = \frac{B(t,T_1)}{B(t,T_2)} \quad \text{i.e.} \quad L = -\frac{B(t,T_2) - B(t,T_1)}{B(t,T_2)(T_2 - T_1)}$$

Note that the rate L is determined by quantities known at time t. It should also be obvious that if a bank offers to enter long and short FRA's with a simple forward rate that is not equal to L, then you will be able to arbitrage them: If the bank offers a rate $F > L$, for example, it is offering too much interest, so you should lend them money. Hence enter into a short FRA, giving you the right to receive a rate of $F\%$ on sa deposit made at time T_1. At the same time buy a T_1–bond and short $\frac{B(t,T_1)}{B(t,T_2)}$–many T_2–bonds. At time T_1, you will receive \$1.00 from the T_1–bond, which you can deposit at a rate of $F\%$. At time T_2, you will have $1 + F(T_2 - T_1)$ in your bank account, but you must pay $\frac{B(t,T_1)}{B(t,T_2)} = 1 + L(T_2 - T_1)$ to honour the T_2–bonds that you are short. Hence at time t, you know that you will have made riskless profit of $(F - L)(T_2 - T_1)$ at time T_2, requiring no outlay of capital. This is clearly arbitrage. If the rate F offered by the bank is $< L$, you can make riskless profits by doing the opposite. Hence in an arbitrage–free environment, $F = L$.

In the above, we dealt with simple forward rates. One can also have discretely or continuously compounded forward rates. For example, the continuously compounded forward rate R should obviously satisfy

$$e^{R(T_2 - T_1)} = \frac{B(t,T_1)}{B(t,T_2)} \quad \text{i.e.} \quad R = -\ln \frac{B(t,T_2) - B(t,T_1)}{T_2 - T_1}$$

It is, in fact, quite easy to see what forward rates should be, using a slightly different argument. Let $F_{m\times n}$ be the (annually compounded) forward rate starting in m years' time and ending in n years' time. Similarly, let $R_{n\times m}$ be the corresponding continuously compounded forward rate. Note that $F_{0\times n}$ and $R_{0\times n}$ are just the n–year rates starting now (at $t = 0$), i.e. they are the n–year spot rates. Suppose that we want to deposit \$1.00 in the bank for 3 years. Then there are a number of strategies we can follow. Here are some:

- Deposit \$1.00 for 3 years. Receive $(1 + F_{0\times 3})^3 = e^{3R_{0\times 3}}$.
- Deposit \$1.00 for 1 year. Also enter an FRA for the period from 1 to 3 years. Receive

$$(1 + F_{0\times 1})(1 + F_{1\times 3})^2 = e^{R_{0\times 1} + 2R_{1\times 3}}$$

- Deposit \$1.00 for 1 year. Also enter an FRA for the period from 1 to 2 years, and another FRA for the period from 2 to 3 years. Receive

$$(1 + F_{0\times1})(1 + F_{1\times2})(1 + F_{2\times3}) = e^{R_{0\times1}+R_{1\times2}+R_{2\times3}}$$

(These are not the only combinations possible.) By arbitrage arguments, these should all amount to the same thing. Thus

$$e^{R_{0\times T_2}T_2} = e^{R_{0\times T_1}T_1+R_{T_1\times T_2}(T_2-T_1)}$$
$$\Longrightarrow R_{T_1\times T_2} = \frac{R_{0\times T_2}T_2 - R_{0\times T_1}T_2}{T_2 - T_1}$$

This gives the T_1–to–T_2 forward rate $R_{T_1\times T_2}$ in terms of the spot rates $R_{0\times T_1}$ and $R_{0\times T_2}$. It is clear that if $R_{0\times T_2} \geq R_{0\times T_1}$, then $R_{T_1\times T_2} \geq R_{0\times T_2}$. Thus if the yield curve of zero rate is upward sloping, then the forward rate curve lies above the yield curve. The opposite is true if the yield curve is decreasing. For example (moving to annually compounded rates), if $F_{0\times1} \geq F_{0\times2} \geq F_{0\times3}\ldots$ is the downward sloping term structure of spot rates, then (with $m \leq n$)

$$(1 + F_{0\times n})^n = (1 + F_{0\times m})^m(1 + F_{m\times n})^{n-m}$$

so that $F_{0\times n} \leq F_{0\times m}$ obviously implies $F_{m\times n} \leq F_{0\times n}$. Thus the from–time–$m$ forward rate curve lies below the spot curve.

We can obviously plot a from–time–t forward rate curve for every time t. When we speak of "the" forward rate curve, however, we generally mean the curve $F(t) = F_{(t-1)\times t}$, where t is in suitable units (years, half–years, quarters, etc.). The spot curve is, of course, given by $S(t) = F_{0\times t}$. We have just seen that $F_{m\times n} \leq F_{0,n}$ if the spot curve is decreasing. Thus, in particular, $F(n) \leq S(n)$, i.e. "the" forward curve lies below the spot curve if the spot curve is decreasing, and above it if the spot curve is increasing.

Example 5.3.2 Suppose that the prices of zero coupon bonds are as given:

Maturity (Months)	**Price**
12	0.915042
18	0.876083
24	0.839256

We can calculate the forward rates directly from the bond prices using the formula

$$F_{T_1\times T_2} = -\frac{B(0,T_2) - B(0,T_1)}{B(0,T_2)(T_2 - T_1)}$$

which we obtained earlier in the arbitrage argument for deriving forward rates. Thus

$$F^2_{12\times18} = -\frac{0.876083 - 0.915042}{0.876083 \times 0.5} = 8.89\%$$
$$F^1_{12\times24} = -\frac{0.839256 - 0.915042}{0.839256 \times 1} = 9.03\%$$
$$F^2_{18\times24} = -\frac{0.839256 - 0.876083}{0.839256 \times 0.5} = 8.78\%$$

Note that we've taken $F^1_{12\times24}$ to be annually compounded, whereas $F^2_{12\times18}$ and $F^2_{18\times24}$ are semiannually compounded forward rates, i.e. we take the compounding rate to correspond to the length of the forward interval. Of course, this can be confusing, and care must be taken to specify the compounding frequency — hence the superscripts.
$F^2_{12\times24}$ can be calculated from $1+F^1_{12\times24} = (1+\frac{1}{2}F^2_{12\times24})^2$, so that $F^2_{12\times24} = 8.83\%$.
We could also have obtained $F^2_{18\times24}$ from $F^1_{12\times24}$ and $F^2_{12\times18}$ as follows:

$$\begin{aligned} (1+F^1_{12\times24}) &= (1+\frac{1}{2}F^2_{12\times18})(1+\frac{1}{2}F^2_{18\times24}) \\ \Rightarrow \quad F^2_{18\times24} &= 2\left(\frac{1+F^1_{12\times24}}{1+\frac{1}{2}F^2_{12\times18}} - 1\right) \end{aligned}$$

The corresponding semiannually compounded spot rates are given by

$$\begin{aligned} F^2_{0\times12} &= 9.08\% \\ F^2_{0\times18} &= 9.02\% \\ F^2_{0\times24} &= 9.00\% \end{aligned}$$

Thus the spot curve is decreasing, and we therefore expect the forward curve to lie below the spot curve. Since $F^2_{12\times18} \le F^2_{0\times18}, F^2_{18\times24} \le F^2_{0\times24}$, it is clear that the (semiannually compounded) forward rate curve lies below the spot curve, as expected.

Let's calculate the annually compounded 2–year par yield c. This solves the equation

$$\begin{aligned} & c(B(0,12)+B(0,24)) + B(0,24) = 1 \\ \Rightarrow \quad & c = 9.16\% \end{aligned}$$

We thus see that $c \ge F^1_{0\times24} \ge F^1_{12\times24}$, i.e. the par yield lies above the spot rate, and the forward rate lies below the spot rate. This is always the case when the yield curve is decreasing.

□

5.3.2 Bootstrapping the Yield Curve

The term structure of riskless zero rates is important for the valuation of all financial securities that have future cashflows, because it contains information about the correct discount factors: One cannot find the present values of securities without knowing the appropriate points on the yield curve. Even though zero coupon bonds are traded for only very short maturities, one can nevertheless back out synthetic riskless zero rates using known yields on coupon bonds and from quoted forward rates.

Example 5.3.3 Given the following data:

Instrument	**Price/Rate**
zero coupon bond maturity 6 months	Price =0.95
FRA for period 6 – 12 months	forward rate $F_{0.5\times1} = 11.21\%$
12% coupon bond face \$100, coupon 12%	Price = 101.12

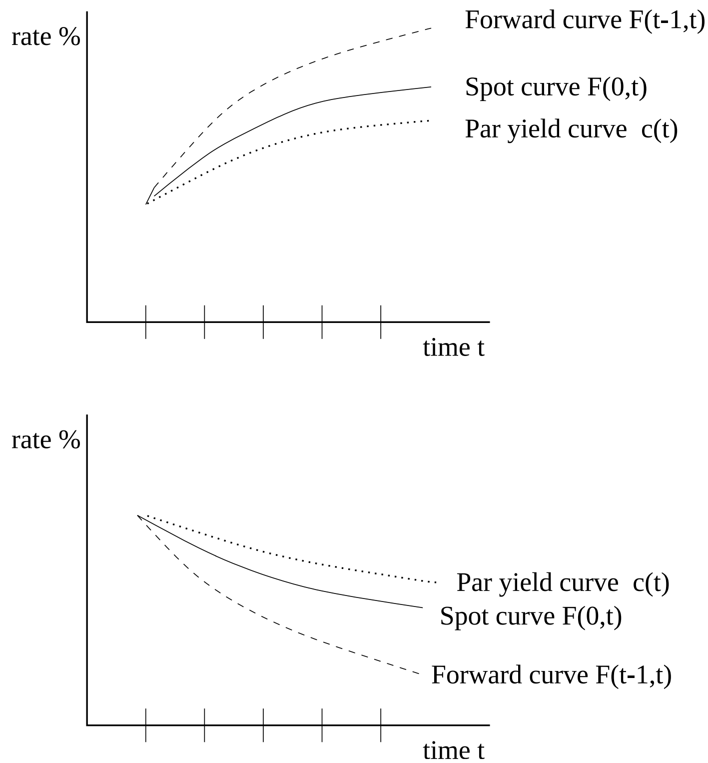

Figure 5.3: Upward- and Downward sloping term structures

We can calculate the continuously compounded spot rates $r(t)$ as follows:

- Since $B(0, 0.5) = 0.95$, we have $r(0.5) = 10.26\%$.
- Next, $e^{r(1.0)} = e^{0.5r(0.5)}(1 + \frac{1}{2}F_{0.5\times 1})$, so $r(1.0) = 10.58\%$.
- Finally, we have

$$\begin{aligned} & 6e^{-0.5r(0.5)} + 6e^{-r(1.0)} + 106e^{-1.5r(1.5)} = 101.12 \\ \Rightarrow \quad & r(1.5) = 10.89\% \end{aligned}$$

□

In practice, we do not have traded instruments with maturities exactly 1.5, 2.0, 2.5 , ... Nevertheless, a fair estimate of zero rates can be obtained by interpolation. For example, suppose we have coupon bonds with maturities 4.8 years and 5.3 years implying zero rates $r(4.8) = 12.24\%, r(5.3) = 12.50\%$, and we want to value a 5–year zero coupon bond. We would then require $r(5.0)$, but this is unavailable. By simple linear interpolation, however, we can find an approximate value:

$$r(5.0) \approx r(4.8) + (5.0 - 4.8)\frac{r(5.3) - r(4.8)}{5.3 - 4.8} = 12.34\%$$

Of course, this is an approximation. Suppose, for example, that we linearly interpolate the semiannually compounded spot curve, and obtain from this the continuously compounded

spot curve. This will give us a curve which differs slightly from one where we linearly interpolate the continuously compounded spot curve (although the curves will agree at times corresponding to the maturity of one of the instruments used).

Using linear interpolation may lead to a yield curve with lots of corners, which is unrealistic. Practitioners may therefore obtained a smoothed yield curve using some other form of interpolation, e.g. by fitting a cubic spline. Consult a book on numerical methods for further information about linear and cubic spline interpolation[5].

5.3.3 Theories of Term Structure of Interest Rates

The yield curve changes over time: It may move up or down, twist, or change shape altogether. An increasing yield curve is regarded as "normal" but an inverted, humped or square–root shaped yield curve may occur as well. How can we explain the shape of the yield curve? There are three major theories: The *expectations* hypothesis, the *liquidity preference* hypothesis and the *market segmentation* (or *preferred habitat*) hypothesis.

The Expectations Hypothesis

The Expectations Hypothesis conjectures that the shape of the yield curve is determined by investors' expectations of future spot rates, i.e. long term rates are determined by expectations of future short term rates. There are several mutually inconsistent versions of the Expectations Hypothesis. The most well–known one asserts that forward rates are equal to the future expected spot rates (Malkiel hypothesis). Suppose that $F_{n\times m}(t)$ is the forward rate, at time t, for the interval $[t+m, t+n]$, so that $F_{0\times n}(t)$ is the n–period spot rate at time t. Then

$$F_{t\times(t+1)}(0) = \mathbb{E}[F_{0\times 1}(t)]$$

i.e. the 1–period forward rate starting at time t is the expected value of the 1–period spot rate at time t. Note that the actual spot rate at time t is unknown, i.e. a random variable. As a result,

$$(1 + F_{0\times n}(0))^n = (1 + F_{0\times 1}(0))(1 + \mathbb{E}[F_{0\times 1}(1)]) \dots (1 + \mathbb{E}[F_{0\times 1}(n-1)])$$

which states that the rate of return on an n–year zero is a geometric average of the expected rates of return of future 1–year zeros. Accordingly, any shape yield curve is posible: If investors expect short–term interest rates to rise in the near future, but to fall again in the long run, a humped yield curve would result. If investors suspect that the government is trying to support the currency by artificially inflating the short term rates, but that this will not last long, a square–root shaped yield curve is a possible outcome.

In practice, the Expectations Hypothesis predicts that investors should be indifferent between investing in long maturity bonds or rolling over a series of short term bonds (in absence of transaction costs).

Liquidity Preference Hypothesis

Borrowers prefer to borrow long term to lock in the interest rate and to ensure the continued availability of capital. Lenders, on the other hand, prefer to lend over shorter periods: The

[5] *Numerical Recipes in C* by Press, Teukolsky, Vetterling and Flannery is a comprehensive cookbook which I highly recommend.

further one tries to look into the future, the more uncertain it is, and risk averse investors would prefer to lend over a period of short terms. This preserves liquidity, in that short term lenders can gain access to their capital at relatively short notice, should future circumstances demand it. Risk averse investors therefore demand a *liquidity premium* in order to be induced to lend over the long term, and thus forward rates should be greater than expected future spot rates. This is in direct conflict with the Expectations Hypothesis, which asserts that the liquidity premium is zero.

In practice, the Liquidity Preference Hypothesis predicts that for bond portfolios the expected return form a a buy and hold strategy should be greater than the expected return of a roll over strategy, even in the absence of transaction costs.

By itself, the Liquidity Preference Hypothesis implies an upward sloping yield curve. Since this is generally observed, it has attractive explanatory power. However, since yield curves are sometimes inverted as well, it is clearly not the whole truth.

Market Segmentation Hypothesis

This asserts that different types of investors have different requirements, and therefore tend to be active in different markets. Hence there need be no relationship short–, medium–, and long term rates. Institutions/firms will attempt to match the maturities of their assets with that of their liabilities in order to reduce interest rate risk. Depending on the nature of their assets/produce, they will therefore borrow either short term or long term. Thus the bond market is segmented into different maturities.

5.4 Risks Associated with Fixed Income Securities

We remarked earlier that governments are amongst the most prolific borrowers, and that government debt (in its own currency) is regarded as "risk–free". What is meant by this is that government debt is free from *default risk*, the risk that the loans will not be repaid. We shall discuss credit risk and credit rating systems in the next section. But there are a host of other kinds of risks associated with investing in bonds, even government bonds, and we briefly describe these here.

- **Interest Rate Risk:** This is *by far the most important* type of risk associated with fixed income securities. As interest rates increase, the value of a fixed income security will decrease, because cashflows must be discounted at higher rates. An investor who needs to sell a bond may therefore suffer severe capital losses if rates increase. Of course, as interest rates decrease, the value of a fixed income security will increase. This may also not be a good thing if an investor intends to hold a bond to maturity, because the interim cashflows (coupons) can be reinvested at lower rates (*reinvestment risk*). Later in this chapter we shall thoroughly discuss the most commonly used measure of interest rate risk, namely *duration*.

- **Inflation Risk:** This refers to the changing value of the purchasing power of cash flows from the security because of inflation. (e.g. if an investor buys a 10-year 6%–coupon bond, but inflation increases to 8%, then the real value of the coupon has declined).

- **Liquidity Risk:** This is the risk that an investor will have to sell an asset below its "true" value (as indicated by recent trades). A security is said to be liquid if it is

relatively easy to buy and sell in large quantities without affecting the price too much. If there is little demand for a particular security, it may be necessary to discount it heavily before a sale can be made. Liquid assets are therefore characterized by tight bid–ask spreads as well as depth (which measures the volume of trade possible without affecting prices). We mentioned earlier that on–the–run securities are generally very liquid. Thus it may be much easier to buy or sell a 30-year Treasury bond issued yesterday, than to buy or sell a 30–year Treasury bond issued 6 months ago, even though there is but little difference between the two instruments. Liquidity risk will even affect investors who plan to hold securities to maturity, if their portfolios need to be marked to market.

- **Political or Legal Risk:** This is the risk of loss due to a political event, such as a change in government: For example, after a civil war a new regime may not feel obliged to honour the debt of the regime it has overthrown. It also refers to losses that are due to a change in legislation, such as new tax laws, or a deliberate currency devaluation.
- **Event Risk:** This is the risk of loss due to such things as a natural or industrial accident, or a corporate takeover/restructuring, new tax laws, changes in economic policy, wars, deliberate currency devaluations, etc.
- **Call Risk:** This is the risk that bond issuers exercise an embedded option.
- **Exchange Rate Risk:** A German investor dealing in dollar–denominated Eurobonds is clearly subject to the risk that the Euro/dollar exchange rate will move in an unfavourable direction.
- **Sector Risk:** Bonds in different sectors (industrials, utilities, mortgages, financials, government) will respond differently to market/economic changes. For example, in a stock market crash there is often a *flight to quality* where investors liquidate positions in corporate bonds and buy government bonds instead. Thus corporate bond prices will decrease drastically, whereas government bond prices may actually increase.

5.5 Credit Risk and Credit Rating Systems

Credit risk or *default risk* is simply the possibility of loss because a counterparty (e.g. a borrower) fails to fulfill certain contractual obligations (e.g. repay its debt on time). When this happens, the counterparty is said to be in default, and bankruptcy proceedings will generally ensue. Credit risk affects virtually every financial contract, but we shall only concern ourselves with defaultable debt, e.g. corporate bonds.

Rating agencies classify bonds according to creditworthiness. Standard & Poor's (S&P), Moody's and Fitch are three of the biggest international rating agencies. It is important to note that a credit rating is *not* investment advice: A consol paying a 0% coupon will no doubt be rated AAA (the highest rating),but it's a lousy investment. Moody's describes a credit rating as " an opinion on the future ability and legal obligation of an issuer to make the timely payments of principal and interest on a specific fixed–income security." A credit ratings may apply to the issuer as a whole, or to a specific issue (e.g. debt collateralized by assets held in trust might have a higher credit rating than the rest of a firm's debt).

A rating is assigned after a thorough analysis of the firm. The firms financial reports and accounting ratios are closely scrutinised, and attention is also given to the structure of

the company, management quality, the nature of a firm's competition, as well as sector– and macroeconomic trends. Both the *probability of default* and the *loss given default* are estimated, using actuarial techniques which rely on historical data.

S&P	Fitch	Moody's	Meaning
		Investment Grade	
AAA	AAA	Aaa	Highest grade, subject to lowest credit risk
AA	AA	Aa	High grade (only slightly riskier than AAA)
A	A	A	Upper–medium grade
BBB	BBB	Baa	Medium grade (adequately protected, but adverse business conditions may impair ability to repay)
		Speculative Grade	
BB	BB	Ba	Lower medium grade (faces major exposure to financial/economic conditions)
B	B	B	Highly speculative
CCC	CCC	Caa	Currently vulnerable to non–payment Default is a real possibility
CC	CC	Ca	Highly vulnerable to non-payment. Default probable.
C	C	C	Default seems imminent. Bankruptcy petition may have been filed.
D			In default

The meaning of S&P,, Fitch and Moody's ratings for long–term debt.

The ratings categories in the table are quite crude, and can be made finer by the use of modifiers with obvious meanings: S&P and Fitch may append a + or − to their rating (e.g. BB+), whereas Moody's appends a number in the range 1-3 (e.g. Aaa2). Different ratings apply to different forms of debt, e.g. S&P categorizes short term debt as A-1, A-2, A-3, B, C or D, whereas Moody's uses Prime-1, Prime-2 and Prime-3.

Rating	Year						
	1	2	3	4	5	10	15
AAA	0.00	0.00	0.09	0.18	0.28	0.67	079
AA	0.01	0.05	0.09	0.19	0.29	0.81	1.21
A	0.06	0.16	0.29	0.45	0.64	1.76	2.61
BBB	0.23	0.65	1.13	1.75	2.38	4.89	7.20
BB	1.00	2.93	5.19	7.36	9.30	16.24	19.43
B	4.57	10.06	14.72	18.39	21.08	28.74	33.26

Cumulative default rates (%) for S&P ratings, 1981–2007.

Rating	Year						
	1	2	3	4	5	10	15
Aaa	0.00	0.00	0.02	0.08	0.16	0.90	1.45
Aa	0.06	0.18	0.29	0.45	0.70	2.29	4.25
A	0.07	0.24	0.50	0.81	1.12	2.90	4.92
Baa	0.29	0.85	1.56	2.34	3.14	7.06	10.44
Ba	1.34	3.20	5.32	7.49	9.59	18.44	25.51
B	4.05	8.79	13.49	17.72	21.43	33.93	41.40

Cumulative default rates (%) for Moody's ratings, 1920–2007.
Source: *FRM Handbook 2011*, by P. Jorion.

The ratings in the tables above are historical ratings, e.g. 2.62% of Baa rated firms had defaulted by their fifth year.

As can be seen from the above tables, credit ratings from different agencies are not directly interchangeable. It should therefore be clear that credit ratings given by different companies do not correspond exactly[6]. A 1993 study showed, for example, that just over half of firms rated AAA or Aaa and AA or Aa where rated the same by both S&P and Moody's. For other investment grade debt, only about 40% where rated the same. It is therefore clear that there is a large element of subjectivity in the ratings process.

Note also that a credit rating is not for life: Ratings are reassessed on a regular basis, and new ratings will be assigned when necessary (for example if the fortunes of a firm take a turn for the better or worse).

Example 5.5.1 In 1988 the leveraged buyout (LBO) of RJR Nabisco[7] resulted in its credit rating being dropped from A1 to B3. The $25 billion takeover was financed largely through debt, and after the take–over, the debt/equity ratio was $29.9 billion/$1.2 billion. The yield spread on existing bonds over the treasury rate jumped from 100 basis points to 350 bp., i.e. the value of bonds issued prior to the take–over plummeted. This was, at the time, the largest LBO ever attempted. Other large corporations, previously though immune to an LBO because of their size, were suddenly perceived to be at risk. Their yield spreads increased as well, even though they were not under attack. Here default risk, event risk and sector risk are at work.

□

Because bonds with a lower credit rating are riskier than those with a higher credit rating, investors require a premium in order to be induced to buy low credit quality bonds. This premium is reflected in the yield: Lower credit quality bonds have higher yields than highly rated bonds. "Riskless" government bonds have the lowest yields and "junk" bonds the highest. The difference in yield between a government bond and a similar credit risky government bond is called the *yield spread* or *credit spread*. Bond traders calculate synthetic zero curves for bonds of each credit quality (using the bootstrapping techniques discussed earlier in this chapter).

[6] As it happens, South African government bonds are currently (January 2013) rated "the same" by all three agencies, namely BBB and Baa1 — one notch above *junk*. Egypt's, on the other hand, is rated BB, B- and B1

[7] *Barbarians at the Gate* by Burrough and Helyar is an entertaining account of the RJR Nabisco LBO

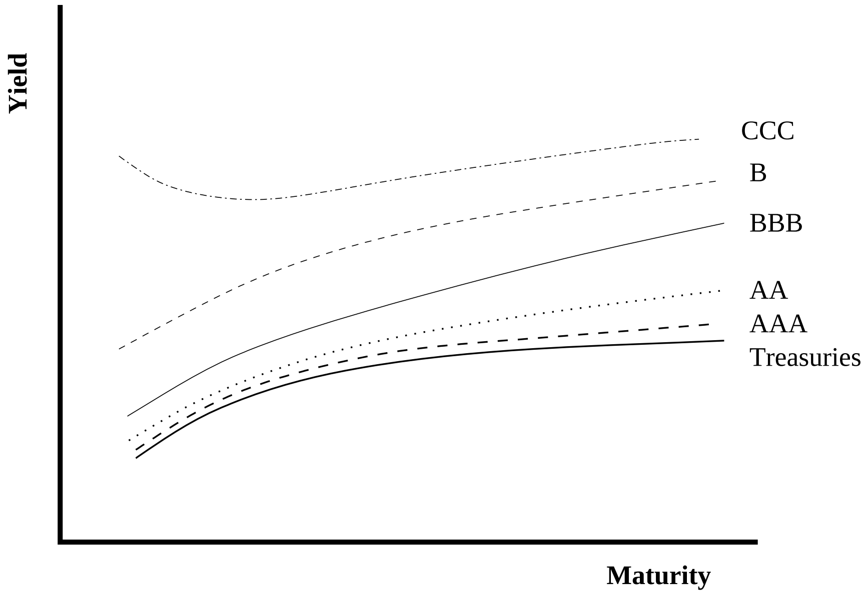

Figure 5.4: Spread curves for different credit qualities

Rating	Spread
Aaa	0.79%
Aa	0.91%
A	1.18%
Baa	1.84%

Average yield spreads on long term investment grade US corporate bonds, 1985 – 1995.
Source: *Credit Risk Valuation* by M. Ammann

However, the yield spread may not be due solely to credit risk. For example, government bonds may be subject to tax breaks, making them more attractive to investors. Yields will therefore reflect this, and the spread between government and corporate debt will be wider than than it would have been without the tax breaks. This must be adjusted for. In practice, traders often use a LIBOR curve as a "riskless" zero curve.

5.6 Traditional Measures of Interest Rate Risk

Since interest rates are generally unpredictable, they form a major component of the risk faced by firms and financial institutions. In this section we look at *duration* and *convexity* as means of obtaining some measure of interest rate risk. These measures will also suggest methods *immunization*, or hedging interest rate risk.

Duration and Convexity of Fixed Coupon Bonds

Portfolio managers are not really worried that interest rates go up: What worries them is that the price of a bond might go down. Of course, these are the same thing. The rate of change of a quantity is measured by its mathematical derivative. For a bond, the interest rate risk can therefore be measured by the derivative of the bond price with respect to the prevailing rate, i.e. the yield. If the bond price is given as a function of the yield, e.g. by

$$P = \sum_{t=1}^{T} \frac{c_t}{(1+y)^t} = P(y)$$

for a T–period bond with cashflows c_t at time t, then if the yield changes by Δy, the bond price changes by

$$\Delta P = P'(y)\Delta y + \frac{1}{2}P''(y)(\Delta y)^2 + \dots$$

where the righthand side is the usual Taylor series. Thus if the yield changes by a small amount Δy, the bond price will change by a small amount

$$\Delta P \approx \frac{dP}{dy}\Delta y$$

which will be negative if Δy is positive (and vice versa). This approximation will quite accurate if $P''(y)$ is close to zero, and less so if $|P''(y))|$ is quite large. The *dollar duration* of a bond is just $-P'(y) = -\frac{dP}{dy}$. The negative sign is placed in the definition to make dollar duration positive: Price is a decreasing function of yield, so $\frac{dP}{dy}$ is negative. The *dollar convexity* of a bond is defined to be its second derivative, $P''(y) = \frac{d^2P}{dy^2}$. Another measure that is occassionally encountered is the *dollar value of a basis point*, denoted DVBP or DV01. It is the amount by which the price of a bond changes if the yield changes by one basis point (0.01%) It is generally more useful to compare relative changes in price, rather than absolute ones. We therefore define the *modified duration* D^* and *convexity* of a bond with yield y by

$$D^* = -\frac{1}{P}\frac{dP}{dy} \qquad C = \frac{1}{P}\frac{d^2P}{dy^2}$$

so that

$$\boxed{\frac{\Delta P}{P} \approx -D^*\Delta y + \frac{1}{2}C\Delta y^2}$$

Modified duration determines the first order (linear) change in the bond price, and convexity takes care of second order effects. Here y is compounded as frequently as is *natural* for this particular type of instrument, e.g. semiannually for a bond that pays a semiannual coupon.

If we take the yield y_c to be compounded *continuously*, then the quantity

$$D = -\frac{1}{P}\frac{dP}{dy_c}$$

is called the *Macaulay duration* (or just "duration") of the instrument. Note that for a fixed coupon T–year bond with face value F a coupon c payable n–times annually, we have

$$P = \left[\sum_{t=1}^{nT} \frac{C_t}{(1+y/n)^t}\right]$$

where $C_t = cF/n$ for $t < nT$, and $C_{nT} = cF/n + F$. Hence the modified duration of the bond is

$$D^* = \frac{1}{nP}\left[\sum_{t=1}^{nT} \frac{tC_t}{(1+y/n)^{t+1}}\right]$$

Here y is the n–times annually compounded yield. On the other hand, the continuously compounded yield y_c satisfies $(1+y/n) = e^{y_c/n}$, so that $y = n(e^{y_c/n} - 1)$. By the chain rule we have

$$\frac{d}{dy_c} = \frac{dy}{dy_c}\frac{d}{dy} = e^{y_c/n}\frac{d}{dy} = (1+y/n)\frac{d}{dy}$$

The Macaulay duration of the bond is therefore

$$D = \frac{1}{nP}\left[\sum_{t=1}^{nT} \frac{\frac{t}{n}C_t}{(1+y/n)^{t}}\right] \tag{*}$$

i.e.

$$D^* = \frac{D}{(1+\frac{y}{n})} \tag{**}$$

In particular, in the continuous compounding case, where $n \to \infty$, we have $D = D^*$.

Also note that

$$C = \frac{1}{n^2(1+y/n)^2P}\left[\sum_{t=1}^{nT} \frac{t(t+1)C_t}{(1+y/n)^{t}}\right]$$

Thus for a fixed coupon bond, duration and convexity are both positive quantities[8].

In textbooks, the Macaulay duration of a bond is often defined by (*), and the modified duration is then defined by (**). I regard this sequence of definitions as more natural from a mathematical point of view.

Macaulay duration was introduced in 1937 by Frederick R. Macaulay, in an influential study of railroad bonds. Macaulay defined the duration of a stream of payments $C_1, \ldots, C_N$ at times $t_1, \ldots, t_N$ by

$$D = \frac{\sum_{i=1}^{N} t_i\overline{\mathrm{PV}}(C_i)}{\sum_{i=1}^{N}\overline{\mathrm{PV}}(C_i)}$$

where $\overline{\mathrm{PV}}(C_i)$ refers to the "present value" of the cashflow C_i at time t, but using the yield as discount rate, and *not* the prevailing spot rate. For fixed coupon bonds, this yields precisely equation $(*)$ above (with $t_i = \frac{i}{n}$). From a financial point of view, however, we see that the Macaulay duration is a kind of weighted average of *times of payments*, weighted by magnitude of payments — hence the use of the term *duration*! If $P = \sum_{i=1}^{N}\overline{\mathrm{PV}}(C_i)$ is the price of the stream of cashflows $C_1, \ldots, C_N$ and if we define $W_i = \frac{\overline{\mathrm{PV}}(C_i)}{P}$, then we see that

$$D = \sum_{i=1}^{N} t_i w_i \qquad \text{where} \qquad \sum_{i=1}^{N} w_i = 1 \tag{‡}$$

Example 5.6.1 If P is a zero coupon bond with face value F, and maturity T

$$D = T\overline{\mathrm{PV}}F/P = T$$

since by definition $P = \overline{\mathrm{PV}}(F)$. Thus the Macaulay duration of a zero coupon bond is precsiely its time to maturity.

[8]We shall soon see examples of securities whose duration and convexity may be negative.

□

Observe that Macaulay duration is measure in units time (years). Modified duration is measured in units of time also, and convexity in units of times squared (years^2). The Macaulay duration of a zero coupon bond is exactly the time to maturity. For a fixed coupon bond, (†) shows that the Macaulay duration is the *average time*, weighted by the bond's "discounted" cashflows as a proportion of the bond price. Thus duration can be thought of as a measure of the average time that a bond is outstanding (instead of maturity). Because a fixed coupon bond pays some of its cashflows before maturity, its duration is less than maturity. Indeed the greater the coupon, the greater the proportion paid off before maturity, i.e. the smaller the duration. You can (and should) check these assertions yourself using (†) or (‡).

Nevertheless, you should think of duration primarily as a *measure of interest rate risk*. Thus the appropriate measure is *modified duration*, not Macaulay duration (because $dP \approx -D^*P\ dy$). In practice, the difference between modified– and Macaulay duration is small, however.

The idea that duration is a kind of average time to maturity works well for fixed income instruments whose cashflows are determined in advance. However, as we shall see soon, *this intuition breaks down badly* for instruments whose future cashflows are uncertain, e.g. FRN's and inverse floaters.

Example 5.6.2 A bond paying a semiannual coupon is trading at par = 100 with a yield of 8%. Suppose that if the yield increases by 1bp, the price of the bond will decrease to 99.95, and that if the yield decreases by 1bp, the price of the bond will increase to 100.04. Thus, taking the average, we have

$$\text{DVBP} = \frac{0.05 + 0.04}{2} = 0.045$$

The modified duration is given approximately by

$$D^* = -\frac{1}{2}\left[\frac{99.95 - 100.00}{+0.0001 \times 100} + \frac{100.04 - 100.00}{-0.0001 \times 100}\right] = \left[\frac{100.04 - 99.95}{0.02}\right] = 4.5 \text{ years}$$

The Macaulay duration is

$$D = D^*(1 + y/2) = 4.5 \times 1.04 = 4.68 \text{ years}$$

□

Example 5.6.3 A 2.5 year bond pays an 8% coupon semiannually. The YTM is 8.12%. We calculate the duration and convexity of this bond. Assuming that the bond has a face value of \$100 (this is not important, because the face value does not affect either duration or convexity), the bond promises the cashflows in the second column. These are then "discounted" using the semiannual yield of 8.12% in the third column:

time t (half years)	Cashflow C_t	$\overline{\text{PV}}(C_t)$	$t\overline{\text{PV}}(C_t)$	$t(t+1)\overline{\text{PV}}(C_t)$
1	4	3.8439	3.8439	7.6879
2	4	3.6940	7.3879	22.1638
3	4	3.5498	10.6495	42.5981
4	4	3.4113	13.6436	68.2265
5	104	85.2343	426.1713	2557.0280
Total		99.7334	461.6980	2697.7044

The bond price is therefore \$99.7334, i.e. slightly below par. We expect this, because the coupon rate is below the yield.

Now recall that Macaulay duration is given by

$$D = \frac{1}{P}\sum_{t=1}^{5} t\overline{\mathrm{PV}}(C_t)$$

Here t is measured in half years, so that

$$\begin{aligned}\frac{461.6980}{99.7334} &= 4.6293 \text{ half years}\\ &= 2.3147 \text{ years}\end{aligned}$$

The modified duration is therefore

$$D^* = \frac{D}{1+\frac{0.0812}{2}} = 2.224 \text{ years}$$

The convexity is given by

$$\begin{aligned}C &= \frac{1}{(1+y/n)^2 P}\sum_t t(t+1)\overline{\mathrm{PV}}(C_t)\\ &= \frac{2697.7044}{99.7334(1.0406)^2} = 24.9796 \ \ (\text{half years})^2\\ &= 6.2449 \ \ (\text{years})^2\end{aligned}$$

□

Example 5.6.4 Consider the same bond as in the previous example. When the yield is 8.12%, its price is 99.73. The modified duration is $D^* = 2.22$, and the convexity is $C = 6.24$. Suppose that the yield changes by +1 bp to 8.13%. The exact new bond price is then 99.71, as you can easily verify using a spreadsheet. Using modified duration, $\frac{\Delta P}{P} \approx -D^*\Delta y$, we get the following approximation for the new bond price:

$$P_{\text{new}} \approx P_{\text{old}} - D^* P_{\text{old}} \times 0.0001 = 99.71$$

If the yield changes by -62 bp to 7.50%, the exact new price is 101.12. The duration approximation, however, is somewhat off: It gives 101.11. If we now incorporate the convexity as well, $\frac{\Delta P}{P} \approx -D^*\Delta y + \frac{1}{2}C\Delta y^2$, the approximate new price is

$$P_{\text{new}} \approx P_{\text{old}} - D^* P_{\text{old}} \times (-0.0062) + \frac{1}{2} C P_{\text{old}} \times (-0.0062)^2 = 101.12$$

If the yield changes by +200 bp, then the exact price, duration approximation and duration–convexity approximation are, respectively: 95.41, 95.30 and 95.42.

This story has the following morals: the duration approximation is good for small changes in yield. For bigger changes in yield, use the convexity as well. However, for changes that are too big, even the duration–convexity approximation is inadequate.

□

There is much geometry in the above example. Recall that the slope of a curve is given by its first derivative. The second derivative is a measure of how curved it is (i.e. the bigger the second derivative at a point, the "benter" the curve at that point). As shown by the graphs below, the duration approximation of the price of a fixed coupon bond is determined by the tangent to the graph of bond price as a function of yield. It is an underestimate. The convexity duration–convexity approximation is better, because it takes into account the curvature of the bond price curve as well as its slope. However, it is a slight overestimate of the true price. This is because for a fixed coupon bond the signs of the Taylor expansion are given by

$$\Delta P = \underset{-}{\frac{dP}{dy}\Delta y} + \underset{+}{\frac{1}{2}\frac{d^2P}{dy^2}\Delta y^2} + \underset{-}{\frac{1}{6}\frac{d^3P}{dy^3}\Delta y^3} + \dots$$

The duration approximation uses the first term, which is negative. The next term is positive, so the duration approximation must have been an underestimate. The third term is negative again, so the duration–convexity approximation, which uses the first and second terms, must have been an overestimate.

How do the coupon rate and maturity of a fixed coupon bond affect its duration? Intuitively, the higher the coupon rate, the greater the proportion of the value of the bond that is paid before maturity. This is immediately obvious from equation (†): Since the Macaulay duration is the cashflow–weighted average of payment times, we expect duration to decrease as the coupon increases (keeping the yield fixed).

The relationship between duration and time to maturity, however, is not that simple. Intuitively, the greater the maturity, the greater the duration, and this is more or less true. We already know that the Macaulay duration of a zero coupon bond is exactly equal to its maturity. Consider now a consol with face value 1, paying annual coupon of c. If the yield is y, then its price is $P = \sum_{t=1}^{\infty} \frac{c}{(1+y)^t} = \frac{c}{y}$, so its modified duration is $D^* = \frac{1}{y}$. The Macaulay duration of a consol is therefore $\frac{1+y}{y}$, independent of coupon. Hence it is possible for a zero coupon bond with finite maturity to have a greater duration than a consol, which has infinite maturity. The duration is therefore not simply an increasing function of maturity. Now a fixed coupon bond with very long maturity is not that different from a consol: Thus we expect the Macaulay duration of a very long term fixed coupon bond with yield y to be approximately equal to $\frac{1+y}{y}$. On the other hand, as the coupon approaches zero, a fixed coupon bond behaves more and more like a zero coupon bond. We thus obtain the graph in Figure 5.6.

We've already pointed out that fixed coupon bonds have positive convexity. The greater the convexity, the more curved the price curve (as a function of yield), i.e. the faster the price curve bends away from a tangent. We now explain why bonds with larger convexity are preferable to those with smaller convexity (assuming equal duration and yield).

In Figure 5.6, bonds A and B have the same yield and duration, but bond B clearly has greater convexity than bond A. If yields increase slightly, the price of A decreases more than the price of B. If yields increase slightly, the price of B increases more than the price of A. Thus no matter whether yields increase or decrease, the owner of bond B is better off than the owner of bond A: The owner of bond B sustains smaller losses if yields increase, and realizes greater gains if yields decrease. "Benter is better".[9]

Generally the market puts a price on convexity, i.e. investors will pay up (accept a lower yield) for greater convexity.

[9]Again, this assumes that the yield curve only experiences parallel shifts.

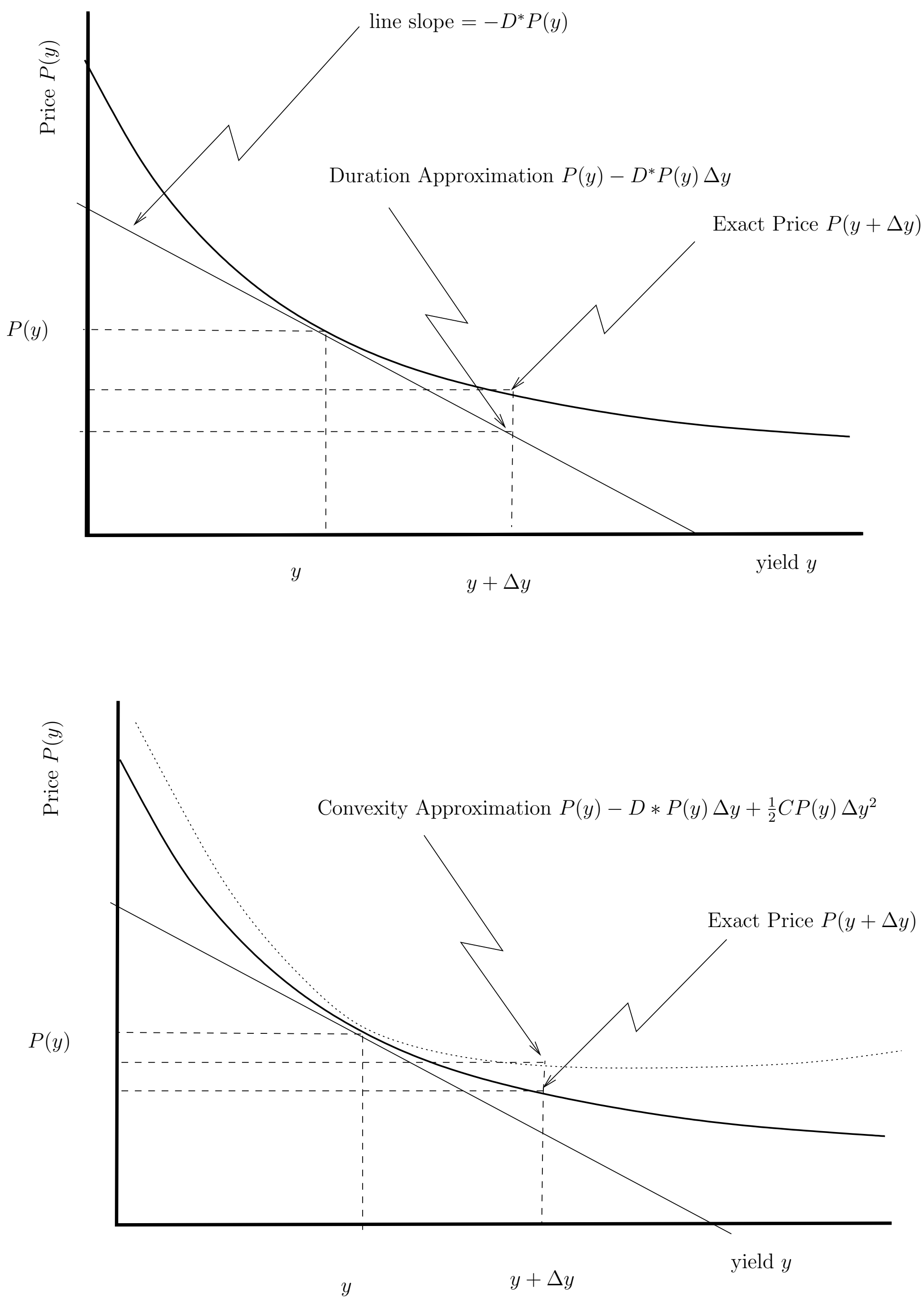

Figure 5.5: Duration and duration–convexity estimates of the price of a fixed coupon bond.

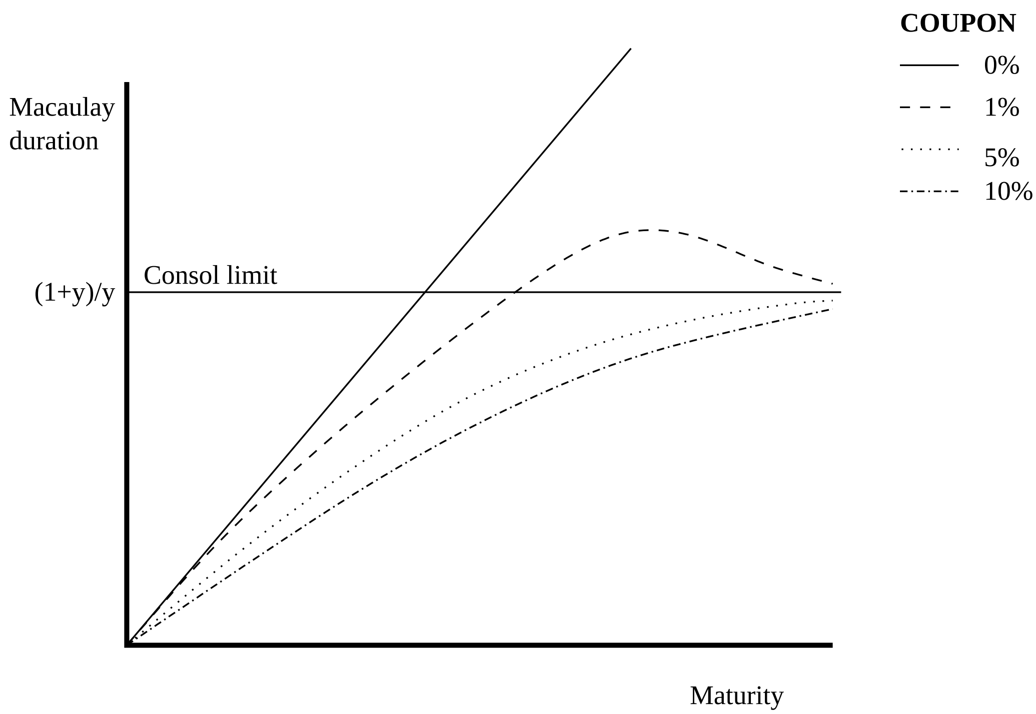

Figure 5.6: The effect of coupon on duration as a function of maturity. All bonds shown have the same yield.

Duration and Convexity of Bond Portfolios

Consider a portfolio P consisting of bonds P_n, so that $P = \sum_n P_n$. Then

$$\Delta P = \sum_n \Delta P_n$$

Up to first order, we therefore have

$$D^* P \Delta y = \sum_n D_n^* P_n \Delta y_n$$

We now make the assumption that each bond yield changes by the same amount, i.e. we assume that the yield curve has undergone a *parallel shift* upwards or downwards. Then $\Delta y_n = \Delta y$. It follows that

$$D^* = \sum_n w_n D_n^* \qquad \text{where } w_n = \frac{P_n}{P} \text{ and } \sum_n w_n = 1$$

Thus the modified duration of the portfolio is simply the weighted average of the durations of each component bond, where w_n is equal to the relative weight of the component P_n as a fraction of the total portfolio value.

The same is true for convexity. We present an argument that would also work for duration: Since differentiation is a linear operation, the dollar convexity of a bond is equal to the sum

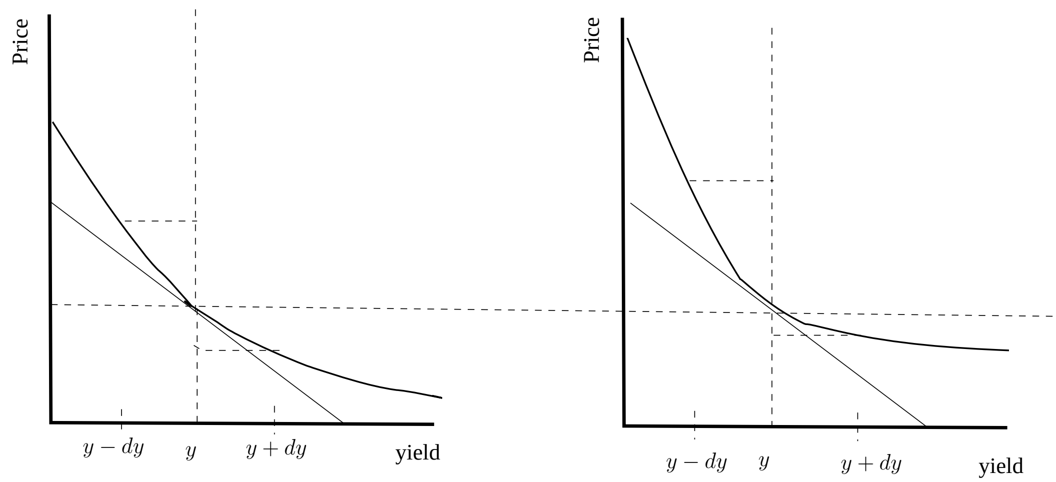

Figure 5.7: Effect of convexity on changes in bond price. Both bonds have the same price and duration when the yield is y.

of the dollar convexities of each component:

$$\frac{d^2P}{dy^2} = \sum_n \frac{d^2P_n}{dy^2}$$

Again, we must assume that all bond prices are functions of the same yield y, or more precisely, that all changes in yields are the same. Hence

$$C = \frac{1}{P}\frac{d^2P}{dy^2} = \sum_n \frac{P_n}{P}\left(\frac{1}{P_n}\frac{d^2P_n}{dy^2}\right) = \sum_n w_n C_n$$

where w_n is the weight of P_n in the portfolio P. Hence the convexity of a portfolio of bonds is equal to the weighted average of the individual portfolio components.

Portfolio managers can hedge interest rate risk by going long on some bonds, and short on others. If a portfolio has the same payoff at some future date, no matter what happens to interest rates, then the portfolio is said to be *immunized.* Suppose for example, that a firm knows that it has to pay an amount Q at time T. Let Q_0 be the present value of this obligation. Thus the firm is essentially short one zero coupon bond with face value Q and maturity T. The simplest hedge is to buy a zero coupon bond with face value Q and maturity T. Suppose, however, that no zero coupon bonds with the required maturity are available. If a pool of bonds is available, then one way to hedge is to choose a bond with the same duration as the obligation, i.e. a bond with duration T. Suppose for example, that an amount Q_0 is invested in a bond P with duration T. If interest rates change, then the new present value of the obligation is approximately $Q_0 - Q_0T\Delta y$, because the duration of the zero coupon bond is T. The bond used to hedge the obligation will have new value equal to $P - D_PP\Delta y$, and since $P = Q_0$ and $D_P = T$, we see that the bond still matches the obligation.

Similarly, suppose a portfolio manager has a portfolio, currently worth Q_0, with a current future value of Q at time T. If interest rates change, the new future value will probably differ from Q. The above argument shows that if the portfolio manager wants to eliminate interest

rate risk, and ensure that Q will be paid off at time T (no matter what happens to interest rates), she should try to ensure that the portfolio has duration T.

Of course, duration is merely a first order measure of interest rate risk. If interest rate moves are too big, second order effects become important as well. In order to hedge, it is therefore desirable that these second order effects don't amount to anything, and this can be achieved by setting the convexity of the portfolio equal to zero as well. (On second thought, positive convexity is a good thing: If the portfolio has duration equal to the investment horizon, and positive convexity, its terminal value will be $\geq Q$.)

Again, it must be stressed that this is true if the yield curve experiences parallel shifts only. With twists in the yield curve, an immunization strategy may fail.

Duration and Convexity for some other Instruments

In the first subsection of this section, we defined the duration and convexity of a fixed coupon bond. We were only concerned with the bond's own yield. In the next subsection, however, we looked at duration and convexity of a portfolio of bonds. This only made sense if we assumed that the all bonds had the same change in yield, i.e. that the yield curve undergoes only parallel shifts. We now generalize one more time to define duration and convexity for any debt instrument. We assume that the value of each instrument is a function of some yield curve y. To be precise, we assume that when interest rates/yields change, the changes are the same for all maturities — i.e. only parallel shifts of the yield curve occur. Thus even though different instruments might have a different y, they all have the same Δy, and hence we can define the modified duration and convexity of any instrument X by

$$D_X^* = -\frac{1}{X}\frac{dX}{dy} \qquad C_X = -\frac{1}{X}\frac{d^2X}{dy^2}$$

Note, however, that some care has to be taken here: If the semiannually compounded zero curve undergoes a parallel shift, then the shift experienced by the continuously compounded zero curve won't be quite parallel. In practice, this seems to make very little difference.

Example 5.6.5 The modified/Macaulay duration of a money market account is zero: When interest rates change instantaneously, the value of the money in the bank is unaffected. Intuitively, a money market account always gives the market–required rate of interest. Hence a money market account has no interest rate risk. The convexity is zero as well.

□

Example 5.6.6 We examine the duration of an FRN F. Suppose that right now it is time t, and that the FRN has future reset dates $T_1, T_2, \ldots$. Also suppose that the face value of the FRN is \$100, and that LIBOR was $x\%$ on the last reset date. At time T_1, the FRN will trade at par, and pay a coupon of x. The value of the FRN at time t is therefore $F = (100 + x)B(t, T_1)$, where $B(t, T_1)$ is a par $= 1$ zero coupon bond with maturity T_1. Clearly, therefore,

$$\begin{aligned} D_F &= -\frac{1}{F}\frac{dF}{dy_c} \\ &= -\frac{1}{(100+x)B(t,T_1)}\frac{dB(t,T_1)}{dy_c}(100+x) \\ &= D_{B(t,T_1)} \\ &= T_1 - t \end{aligned}$$

because the Macaulay duration of a zero coupon bond is just its time to maturity. Hence the Macaulay duration of an FRN is equal to the time to the next rest date.

Intuitively, this is obvious: At any time, an FRN is essentially just a zero coupon bond plus a money market account. The zero coupon bond corresponds has a face value equal to the par value plus the next coupon, and matures at the next reset date. But at the next reset date, the new coupon rate will be set to the prevailing market rate, just like a money market account. Since a money market account has no interest rate risk, all the interest rate risk resides in the zero coupon bond. Hence the duration of the FRN is the same as that of a zero coupon bond.

□

Example 5.6.7 We estimate the duration of an inverse floater. Consider a just issued par \$1 million inverse floater paying 24% - 2LIBOR, when prevailing rates are approximately 8%.

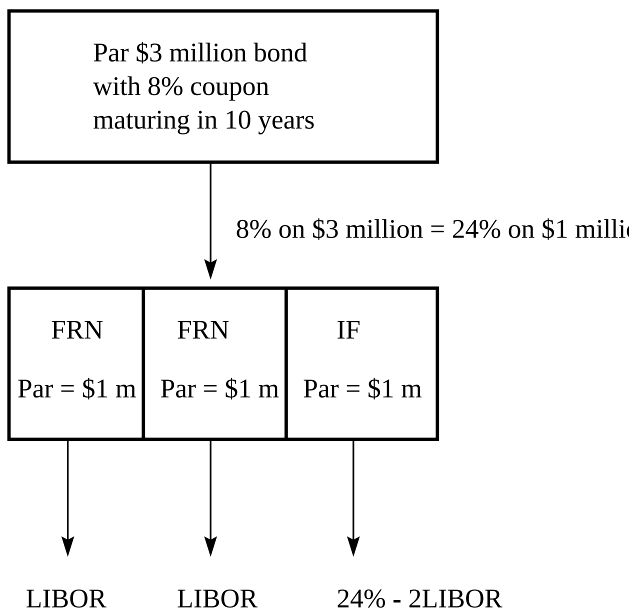

Figure 5.8: IF = Bond + FRN

As shown by the diagram, an inverse floater plus two FRN's make up a fixed coupon bond. Now the duration of the FRN's is very small (compared to 10 years), and we'll ignore it. Moreover, an FRN trades at values close to par (=\$1 million), and the fixed coupon bond trades close to par as well, because its coupon is close to current interest rates. The value of the inverse floater is therefore approximately \$1 million. The duration of the fixed coupon bond is approximately 7 years. Since the duration of a portfolio is the weighted average of its components, we see that

$$D_{\text{bond}} \approx \frac{1}{3} D_{\text{FRN}} + \frac{1}{3} D_{\text{FRN}} + \frac{1}{3} D_{\text{IF}}$$

and conclude that the duration of the inverse floater is about 21 years. This is far greater than its time to maturity.

□

The above examples show that for a debt instrument with uncertain future cashflows, duration may be completely unrelated to maturity. Perhaps an FRN matures in 10 years' time; the duration is the time to the next reset date. The duration of an inverse floater can easily be twice the time to maturity. Thus:

Duration is a measure of **interest rate risk**. It can only be regarded as a cashflow adjusted time to maturity for debt instruments whose future cashflows are *known*.

Example 5.6.8 Estimates of duration and convexity are more difficult to obtain for callable bonds, and require techniques from option pricing theory. We show here, however, that callable bonds can have negative convexity.

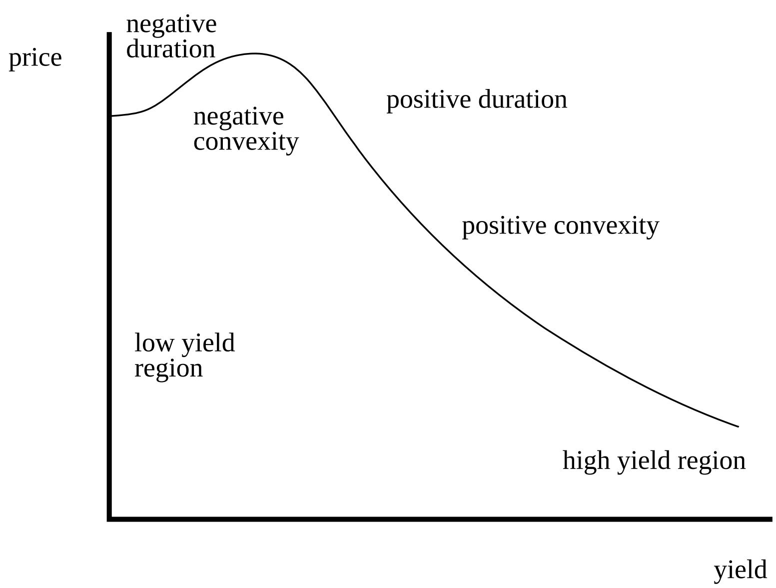

Figure 5.9: Price–Yield Curve for Callable Bonds

Suppose that an investor holds a callable bond. Since the firm that issued the bond has the right to call it back, the investor is short a call option. The issuer will only exercise this option if interest rates become very low, because then it will pay to call back the bond and float a new issue with a lower coupon. In the high yield region, therefore, the bond has a very small chance of being recalled. The call is therefore nearly worthless, and the callable bond behaves just like a regular bond.

In the low yield region, on the other hand, a drop in yield will cause the price of a regular bond to increase, but it will also make exercise of the call more likely, i.e. the call increases in value as well. Thus there is a tug of war between the price changes of the bond and the option. This causes the callable bond price to increase less rapidly when yields decrease than it would have if it was a regular bond, resulting in negative convexity. It is in fact possible

for a callable bond's price to decrease if yields decrease. Thus duration may be negative as well.

□

Example 5.6.9 *Mortgage–Backed Securities* (MBS's) are debt securities collateralized by property. The buyer of an MBS essentially is providing a loan with which a homeowner can pay off a house. The owner of an MBS then receives cash flows determined by the mortgage loan payments of property owners. The homeowner, however, has the opportunity to py back his loan early, e.g. if he decides to move or to refinance. Essentially, the homeowner is long a call option, and since he will tend to exercise this option only when it favours him, this results in a risk, prepayment risk, for the investor in MBS. Moreover, the homeowner also has the option to increase the amount of monthly mortgage payments, shortening the life of the loan — contraction risk. Because the investor in MBS is short a call option, MBS have regions with negative convexity, just as in the previous example.

Since the life of an MBS is uncertain, it is difficult for certain investors, such as pension funds, to invest in them: It is difficult to match the cash flows received from the MBS with the pension that must be paid out, for example. One way around this difficulty is the use of **Collateralized Mortgage Obligations** (CMO's).

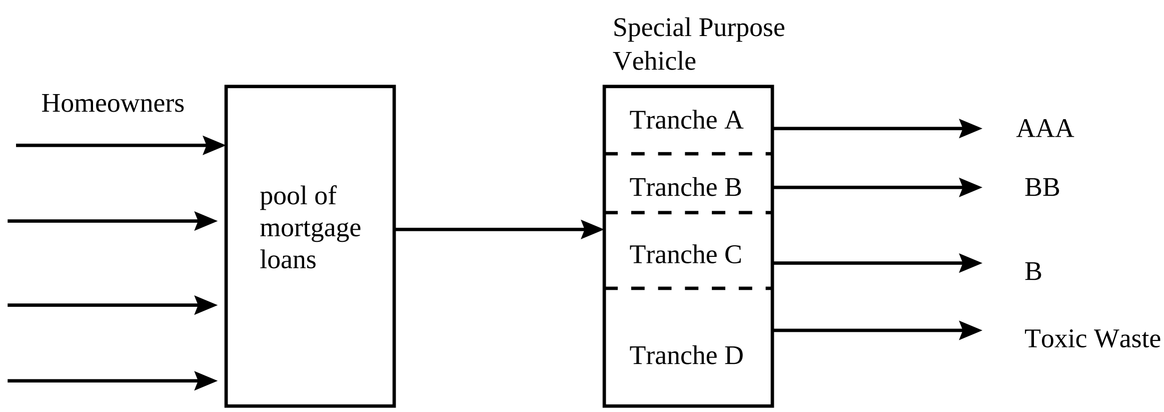

Figure 5.10: Structure of a Collateralized Mortgage Obligation

In a CMO, mortgage loans are pooled together. The cash flows from the mortgage loans go into a Special Purpose Vehicle (SPV), a legal entity which issues securities with different characteristics. For example, the benefits of diversification will make it possible to practically guarantee that certain payments will be made, and these will correspond to AAA–rated securities. Other tranches will have less predictable payouts; these have lower credit ratings, but higher yields.

Note that what goes into the SPV must come out, i.e. no new money is created, and the total risk of the SPV is the same as that of the pool of mortgage loans. All that the SPV does is redistribute cashflows and risks: The total amount paid out by the SPV is the same as what is paid in by homeowners. If some tranches are less risky than the collateral, others must be more risky. The top tranche may have a very high credit rating, but the bottom tranche (often called the *toxic waste*) is very risky indeed.

Mortgage loan payments include not only the interest on the loan, but also part of the principal (amortization). Another common way to structure CMO's is into *Interest Only* and *Principal Only* tranches. As you no doubt already guessed, the IO/PO structure separates the interest from the principal. The IO tranche receives only the interst payments, and the PO tranche only the principal.

When interest rates go up, homeowners will pay more interest, and will therefore tend to pay back less of the principal. When interest rates fall, two factor contribute to the increase in the value of the PO tranche: the principal will be paid back earlier, and at lower discount rates. The opposite happens if interest rates go down.

Hence the IO tranche increases in value if interest rates go up: It has negative duration. The PO tranch, on the other hand, has positive duration.

□

For debt securities with cahnging cash flow patterns, duration and convexity may be either intractable or unreliable. Instead, sensitivity is measured using *effective duration* and *effective convexity*:

$$\begin{aligned} D^* &= -\frac{P^+ - P^-}{2P\Delta y} & P &= P(y) \\ &= \text{average of left– and right derivatives} & P^+ &= P(y+\Delta y) \\ C &= \frac{P^+ - 2P + P^-}{P(\Delta y)^2} & P^- &= P(y-\Delta y) \end{aligned}$$

Limitations of Duration and Convexity

As we have stated on a number of occasions already, duration and convexity are effective risk measures only if it is assumed that shifts in the term structure are small and parallel. For example, the formula $\Delta P = -D^* P \Delta y$ relates the price change of a bond to a change in yield/interest rates. For this to be accurate, Δy is required to be small. To measure risk accross different bonds, we must assume that the yield of each bond changes by the same amount, i.e. Δy is the same for all bonds. This means that the yield curve has undergone a parallel shift upwards or downwards.

In practice,, this is generally not the case.

- Short term rates tend to be more volatile than long term rates.
- On–the–run securities are more heavily traded than off the run securitie, leading to higher volatility.
- The long– and short term sections of the yield curve may move in opposite directions.
- The yields of corporate bonds may move in a direction opposite to that of government bonds, e.g. when investors dump corporate bonds during a "scare" and invest in Treasuries instead ("flight to quality").

Nevertheless, several empirical studies have used principal component analysis to show that the movements of the yield curve can generally be decomposed into three independent factors:

(a) A parallel shift of the yield curve, accounting for $\approx 80 - 90\%$ of the total variance.

(b) A twist, where long– and short term rates move in opposite directions. This accounts for $\approx 5 - 10\%$ of total variance.

(c) A “butterfly”, where medium term rates move in directions opposite to long– and short term rates. This generally accounts for about $1 - 2\%$ of total variance.

Since most of the motion of the yield curve is due to a parallel shift, the use of duration and convexity is to some extent justifiable.

Chapter 6

Fixed Income Derivatives

There exist a large number of OTC and exchange traded fixed income derivatives (also called *interest rate derivatives*). These are securities whose value is derived from the value of a fixed income instrument or the level of an interest rate. We have already met one example, namely FRA's, in Chapter 3. In this chapter we introduce swaps, futures, caps and floors. Swaps and futures are linear instruments, and thus relatively simple to price. Caps and floors, however, are options on a prevailing interest rate, and therefore require more advanced techniques from option pricing theory. We shall thus content ourselves with a mere description of interest rate options.

6.1 Valuing Forward Rate Agreements

We briefly recall some facts from Chapter 5: A forward rate agreement (FRA) is an OTC contract between a customer and a bank that gives the customer a guaranteed rate of interest on a specified amount over a specified future period. The guaranteed contract rate is the *forward rate* for that period, and is chosen so that the initial value of the contract is zero. However, the forward rate changes from day to day, whereas the contract rate remains fixed. The difference between the forward rate and the contract rate of an FRA is similar to the difference between the forward price of an asset and the delivery date of a forward contract: On the day that a forward contract is initiated, the delivery price and the forward price are the same. But the forward price changes from day to day, and the delivery price of a particular forward contract remains fixed.

The fact that the forward rate can be different from the contract rate means that a forward contract can have non–zero value after it is initiated. This value is rather easy to calculate.

Example 6.1.1 Suppose that we are given the following information about the yield curve (continuously compounded):

Maturity (years)	Spot Rate (CC)
0.25	8.09%
0.50	8.35%
0.75	8.51%
1.00	8.62%

At $t = 0$, we enter into a 6–month FRA on a notional of \$1 million, starting in 6 months' time. By arbitrage arguments, borrowing \$100 million at the 1–year spot rate should give us the same as borrowing for 6 months at the 6–month spot rate, followed by a 6–month FRA. Thus

$$e^{0.0862} = e^{0.0835/2} e^{R_{6\times 12}(0)/2}$$

where $R_{6\times 12}(0)$ is the $t = 0$ continuously compounded forward rate for the period from 6 to 12 months. Hence $R_{6\times 12}(0) = 8.89\%$. This is the guaranteed rate of our FRA. Generally, the rate would be quoted as a semiannually compounded rate, i.e. 9.09%.

In 3 months' time, the yield curve looks like this:

Maturity (years)	Spot Rate (CC)
0.25	8.55%
0.50	8.89%
0.75	9.12%
1.00	9.29%

Our FRA is now a 6–month FRA starting in 3 months' time. To have zero value, our the contract would have to guarantee a rate of

$$R_{3\times 9}(0.25) = \frac{0.75 \times 0.0912 - 0.25 \times 0.0855}{0.75 - 0.25} = 9.41\%$$

which would be 9.63% semiannually compounded.

However, the contract rate of our FRA is still 8.89%. Thus the interest that we will have to pay over the 6–month life of the FRA is less than what the market demands, i.e. the FRA now has positive value. This value is simply

$$\text{Value of FRA} = \$1 \text{ million}(e^{0.0941/2} - e^{0.0889/2})e^{-0.0912\times 0.75} = \$2\ 542$$

(where we have discounted from the maturity of the FRA in 9 months' time.)

Alternatively, using the semiannually compounded rates, the value of our FRA is

$$\text{Value of FRA} = \$1 \text{ million}(9.63\% - 9.09\%) \times \frac{1}{2} \times e^{-0.0912\times 0.75} = \$2\ 542$$

□

Note that a party which is *long* an FRA will borrow at a fixed rate, whereas a party *short* an FRA will lend at a fixed rate. Thus a long FRA protects against rising interest rates, whereas a short FRA protects against falling rates.

We conclude this section with a few word about the duration of an FRA. If we look at the mathematical definition of the duration of an instrument with value P,

$$\text{Duration} = -\frac{1}{P}\frac{dP}{dy}$$

you will note that this doesn't make sense for instruments that initially have zero value, for example FRA's and swaps. For an FRA, the duration is clearly $-\infty$ at the start of the contract. Nevertheless, you will find a number of texts in which the durations of FRA's and swaps are mentioned, and even calculated, and it is not always clear what this means. In this

case the authors seem to take the notional amount as the value of the instrument, i.e. they use something like the following as definition of the "duration" of a contract:

$$\text{Duration} = \pm \frac{1}{\text{Notional Amount}} \frac{d\text{Value}}{d \text{ Rates}}$$

For example, the "duration" of a 6×12–FRA on a notional of \$1 000 000 can be calculated as follows: If interest rates go up by Δy (semiannual compounding), the value of the FRA at its expiry increases by

$$\$1 \text{ million} \times \Delta y \times (12 \text{ months} - 6 \text{ months})$$

The duration is therefore just (12 months − 6 months) = 0.5 years. It can thus be seen that the "duration" of an FRA starting at time T_1 and ending at time T_2 is just $T_2 - T_1$. In this way it is possible to use FRA's to roughly hedge interest rate risk, an example of which is given in Section 6.3.2.

6.2 Interest Rate Swaps

6.2.1 The Swap Market

The following definition has been adapted from a number of sources:

> A *swap* is an agreement by two counterparties to *exchange* a predetermined *series of cash flows*, decided by a preagreed formula, over time.

This definition makes almost any non–optional financial contract a swap contract: For example, a forward contract is essentially an ageement to swap the $t = 0$–forward price of an asset for the $t = T$–spot price at a future date T. When the preagreed formula involves interest rates, we have an *interest rate swap*, the topic of this section; when it involves payments in different currencies, a *currency swap*. *Commodity swaps* and *equity swaps* also exist.

Interest rate swaps have payments linked to one or more reference rates. The amount of interest exchanged is based on a predetermined notional amount, but only the interest payments are exchanged, not the notionals. Payments are often *netted*, i.e. only the difference between the required interest rate payments is exchanged. This reduces credit risk.

Example 6.2.1 In 1974, Bankhaus Herstatt went bankrupt. On the day that bankruptcy was declared, Herstatt received payments on foreign exchange trades from a number of counterparties, but defaulted before they made the payments that they owed in return. The risk that a party defaults on one leg of two offsetting and nearly simultaneous payments is called *settlement risk* or *Herstatt risk*. Herstatt risk only lasts for a very short while: between the time that one party has made a payment, and the time that the offsetting payment is received. This could be a couple of hours, because of differences in trading times between different countries. Because of the Herstatt debacle, mechanisms are now in place to ensure that payments are netted: For example, if party A owes party B \$1 million on an interest rate swap, and party B owes party A \$990 000 in return, then under a *netting agreement* only \$10 000 passes from A to B. Without a netting agreement, it is possible for B to pay A \$990 000, and receive nothing in return (due to default). Hence netting agreements can significantly reduce credit risk.

□

The most common type of interest rate swap is the *fixed–for–floating* swap (also called a plain vanilla swap, a coupon swap, or a generic swap). These are OTC instruments in which one party agrees to make payments on a notional amount at a fixed rate in return for floating rate payments. The reference floating rate is most often a LIBOR rate, but may also be a rate implied by Treasury bills, banker's acceptances or CD's. *Basis swaps* have payments from both parties linked to a (different) floating rate.

The first swap contracts were organized by the World Bank in 1981. Since then, the market for swaps has grown tremendously, and they are by far the most common type of derivative traded today, accounting for over 40% of the amounts oustanding in the global derivatives market. According to the December 2002 BIS Quarterly Review, the global derivative markets (OTC and exchange traded) amounted to $\approx$ \$155 trillion outstanding in June 2002. OTC interest rate derivatives amounted to $\approx$ \$90 trillion, and of these, interest rate swaps to $\approx$ \$68 trillion. In 1999, interest rate swaps amounted to only $\approx$ \$44 trillion, which shows that the swap market is still experiencing tremendous growth. For comparison, note that equity linked options (OTC and exchange traded) amounted to a paltry \$3.5 trillion outstanding in 2002.

6.2.2 Fixed–for–Floating Swaps

The best way to explain the mechanics of a swap is via an example. Suppose that corporation A has borrowed from a bank at floating rates, but would prefer to pay a fixed rate. Corporation B, on the other hand has issued a fixed coupon bond, but is of the opinion that shareholder value will be maximized if it transforms its debt to floating rate debt. A and B are made for each other! Consider therefore a three–year swap initiated on 7 February 2002 by the two counterparties, A and B. A agrees to pay B a fixed rate of 10% on a notional of \$100 million, making semiannual payments. In return, B agrees to pay B 6–month LIBOR. If 6–month LIBOR was 9% on 07/02/2002, then A will pay B an amount of \$5 million on 07/08/2002 (being half of 10% of \$100 million), and will receive \$4.5 million in return. LIBOR, meanwhile, is not sitting still, and B's next payment is a random variable. If LIBOR moves according to the 2nd column in the table below, then the cash flows are given by the third and fourth columns:

Date	6 month LIBOR	A pays million US$	B pays million US$	Net to A million US$
07/02/2002	9%			
07/08/2002	10%	5.0	4.5	-0.5
07/02/2003	12%	5.0	5.0	0.0
07/08/2003	11%	5.0	6.0	1.0
07/02/2004	9%	5.0	5.5	0.5
07/08/2004	9%	5.0	4.5	-0.5
07/08/2005	10%	5.0	4.5	-0.5

Here A is the fixed rate payer, and B the floating rate payer. This is an *in–arrears* swap, because B pays A the floating rate at the end of the corresponding period. *In–advance* swaps, where the floating rate payer pays the rate correponding to the next period, also exist. In an in–advance swap, for example, B would have paid \$1.0 million on 07/02/2003 (and not on 07/08/2003).

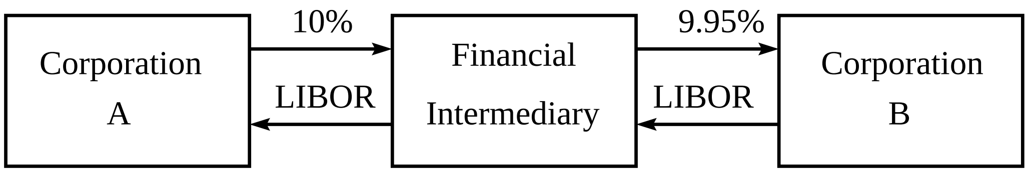

Figure 6.1: A swap agreement brokered by a financial institution.

If payments are *netted*, only the amount in the fifth column is exchanged.

The rate paid by the fixed rate payer is called the *swap rate*, and is chosen so that the *swap initially has zero value*. In this way, a swap is like a forward contract. Just like the forward price of an asset changes from day to day, so does the swap rate. When the swap rate increases from its initial value, A has to pay less interest than the fair market rate. Hence the value of the swap is now positive to A (and negative to B). We'll say more about swap valuation in a later section.

In practice, corporation A and B do not generally get in touch with each other. Instead A and B will each approach a financial intermediary, such as a bank, who will try to match counterparties, and keep a small spread. Since it is unlikely that A and B approach the financial intermediary at the same time, the intermediary will act as a counterparty itself, hedging its position with other interest rate derivatives. Thus generally, A and B are not even aware of each other's existence, and A and B each have a contract with the intermediary, not with each other. If A defaults on its payments, it is the intermediary that suffers, not B. The competition amongst financial institutions for swap contracts is very intense, and the spread pocketed as payment for brokering a swap deal is quite small. In the USA, for example, it generally amounts to about 5 bp (as it does in Figure 4.1).

6.2.3 Why Enter into a Swap?

Why would two institutions want to enter into a swap, if swap contracts are zero sum? One reason is interest rate risk. A corporation servicing floating rate debt is at risk if interest rates go up, whereas a corporation with fixed rate debt faces the risk of interest rates coming down. In order to hedge, it may be desirable to have a mixture of fixed and floating rate debt. *Swaps transform fixed rate debt into floating rate debt*, and vice versa.

Example 6.2.2 Bank A has a portfolio of 5 year corporate loans (it has made) with a face value of \$100 million. Interest on all loans is 10%, paid semiannually. To fund this loan, the bank depends on 6 month CD's, on which it pays 6 month LIBOR + 40 bp (i.e. depositors deposit \$100m which the bank lends to corporations). Thus A faces the risk that the payments it owes the holders of the CD's are not covered by the payments made by the corporation, i.e. it faces the risk that LIBOR $\geq$ 9.6%.

Life insurer B is committed to pay 9% on a guaranteed investment contract, on a notional of \$100 million. B has the opportunity to invest in an FRN paying LIBOR +160 bp semiannually. It will then face the risk that its obligations on the guaranteed investment fund are not offset by the coupons from the FRN, i.e. it faces the risk that LIBOR $\leq$ 7.4%

Thus A and B face opposing risks: A is worried that LIBOR will increase, and B is worried that LIBOR is increase. A and B are made for each other. Moreover, if 7.4% $\leq$ LIBOR $\leq$ 9.6%, then both A and B are making money. There is a spread of 9.6% - 7.4% = 2.2% which

they ought to be able to lock in on at no risk whatsoever[1]. Financial intermediary C brokers a deal whereby it keeps 10 bp and passes 1.05% to each client (so that 1.05% + 1.05% +10 bp = 2.2%).

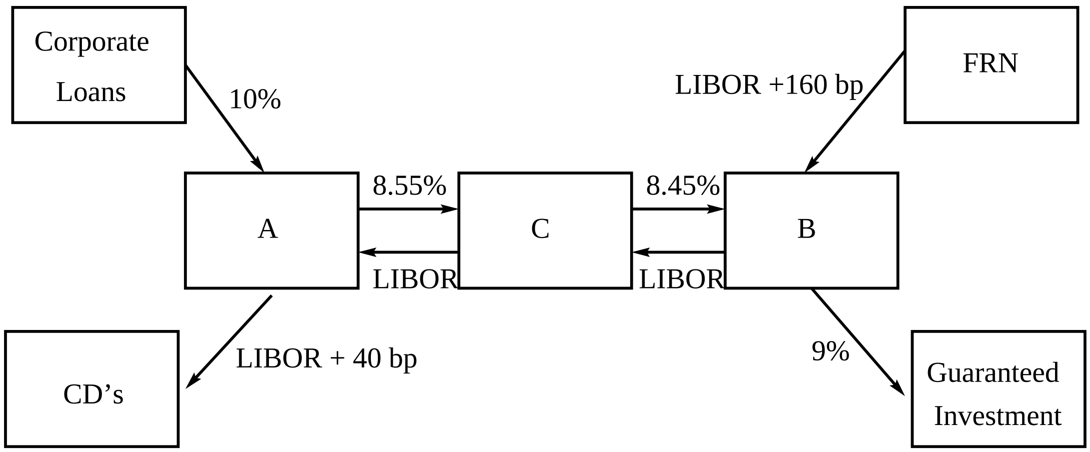

Figure 6.2: Swap deal

For example, A faces the risk that LIBOR will increase. To offset this risk, it should receive LIBOR, and thus pay a fixed rate. What should this fixed rate be? If A is going to receive a spread of 1.05%, then the fixed rate x must be chosen so that $10\% + \text{LIBOR} -$ (LIBOR + 140 bp) $- x\% = 1.05\%$, so that $x = 8.55\%$. C pockets 10 bp, and pays 8.45% to B. B makes a total of LIBOR + 160bp + 8.45% − LIBOR − 9% = 1.05%. Every one makes money.

Is this deal realistic? Probably not very, because every one makes money. If A is paying LIBOR +40 bp and the FRN is paying LIBOR + 160 bp, then the FRN probably carries a large amount of default risk, which we have ignored. B is probably getting a raw deal here: If LIBOR goes up, the chances that the FRN defaults will increase at the same time that B is reponsible for higher interest payments.

□

The Comparative Advantage Argument

The usual explanation for the popularity of swaps goes along the following lines: Some companies have a comparative advantage when borrowing in fixed rate markets, wheras others have an advantage in floating rate markets. A firm that has an advantage borrowing at fixed rates, but wants a floating rate loan, may find it advantageous to enter into a swap.

Example 6.2.3 Consider two companies X and Y, who have been offered the following rates on a $10 million 5–year loan:

	Fixed	Floating
X can borrow at	8.0%	LIBOR +10 bp
Y can borrow at	8.8%	LIBOR + 50 bp

[1] *Market risk*, that is, not default risk.

X wants a floating rate loan, and Y wants a fixed rate loan. Note that X has been offered lower rates in both the fixed– and the floating rate market; it has an *absolute advantage*, presumably because it is less credit risky. However, the spread of X over Y in the fixed rate market is 80 bp, whereas in the floating rate market it is only 40 bp. Thus X is doing comparatively better in the fixed rate market, whereas Y is doing comparatively better in the floating rate market. It will be to their advantage to enter a swap. The difference in spreads is 80 - 40 = 40 bp. Assuming that this is shared equally between X and Y, with 10 bp going to a financial intermediary, X and Y should both be able to make 15 bp.

To work out the mechanics, consider the payments made by Y. Though Y wants a fixed rate loan, it takes out a floating rate loan and enters a swap. Similarly, X takes a fixed rate loan. Y will now pay a fixed rate $x\%$ to the financial intermediary, pay LIBOR+50 bp, and receive LIBOR, a total fixed rate of $x + 50$ bp. Y would have had to pay 8.80% in the fixed rate market, but since it gains 15 bp, it only pays 8.65%, i.e. $x + 50 \text{ bp} = 8.65\%$. Thus Y pays 8.15% to the intermediary. The intermediary pockets 10 bp, and pays 8.05% to A. A ends up paying a net LIBOR - 5 bp, instead of LIBOR + 10 bp.

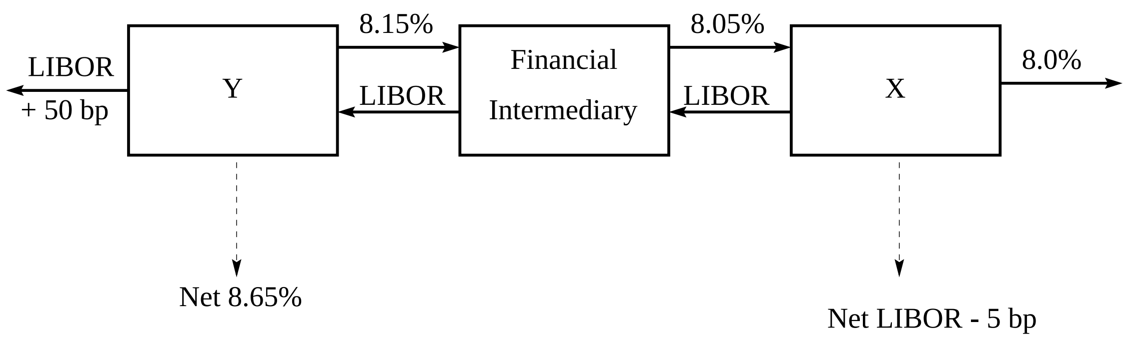

Figure 6.3: Swap deal

□

6.2.4 Swap Valuation

A swap can be decomposed into two ways:

- A receive–fixed swap is equivalent to a portfolio which has a long position in a fixed–coupon bond and short position in a floating rate note.
- A receive fixed swap can also be seen as a portfolio of short FRA's, one for each exchange date, and each with the same "forward rate".

Of course, a pay–fixed swap is just a short receive–fixed swap, and therefore corresponds to the opposite portfolios.

Example 6.2.4 Consider a 3–year receive–fixed swap, where we receive a fixed coupon of 5.5% in exchange for LIBOR on a notional of $100 million, with payments made annually. Assume that the term structure of zero coupon bonds is given by the 2nd column of the table below. then the annually compounded spot– and forward rates are given by the third and fourth columns:

T	$B(0,T)$	Spot Rate	Forward Rate
1	0.9615	4.00%	4.00%
2	0.9136	4.62%	5.24%
3	0.8591	5.19%	6.34%

The cash flows of this swap are the same as for a portfolio consisting of long fixed coupon bond with a face value of \$100 million, paying an annual coupon of 5.5%, and a short FRN with a face value of \$100 million, paying annual LIBOR. The value of the fixed coupon bond is

$$5.5 \times 0.9615 + 5.5 \times 0.9136 + 105.5 \times 0.8591 = \$100.95 \text{ million}$$

The FRN trades at par, and so is worth \$100 million. The value of the swap is therefore

$$\begin{aligned}\text{Receive–fixed swap} &= \text{Fixed coupon bond - FRN}\\ &= 100.95 - 100\\ &= \$0.95 \text{ million}\end{aligned}$$

We could also regard the swap as a portfolio of short FRA's, where we agree to pay LIBOR in return for a "forward rate" of 5.5%. Thus we must value each FRA. Now the 0×1 forward rate is 4% and we will receive 5.5% at time $t = 1$. Hence the value of the first FRA is

$$\text{Value(FRA 1)} = 100 \text{ million} \times [5.50\% - 4.00\%] \times 0.9615 = \$1.44 \text{ million}$$

Similarly,

$$\text{Value(FRA 2)} = 100 \text{ million} \times [5.50\% - 5.24\%] \times 0.9136 = \$0.23 \text{ million}$$

and

$$\text{Value(FRA 3)} = 100 \text{ million} \times [5.50\% - 6.35\%] \times 0.8591 = -\$0.72 \text{ million}$$

Thus the value of the swap is

$$\begin{aligned}\text{Receive–fixed swap} &= \text{Value(FRA 1)} + \text{Value(FRA 2)} + \text{Value(FRA 3)}\\ &= \$0.95 \text{ million}\end{aligned}$$

□

Even though a swap can be seen as a portfolio of FRS's, swaps are not redundant instruments, because FRA's are not generally quoted beyond two years' maturity, and liquid only over a one year period. The swap market is far more liquid than the forward market. Moreover, a 10–year swap with semiannual payments would require about 20 separate FRA contracts, whereas the swap contract is a single contract: It saves on transaction costs.

The Swap Rate

As we've stated before, the rate paid by the fixed rate payer is called the *swap rate* and is chosen so that *the initial value of a swap contract is zero* (par swap). In this respect, a swap is like a forward contract or an FRA. Since a swap is just a portfolio long a fixed coupon bond and short an FRN, the swap rate c must be chosen so that

$$\text{Value of bond with coupon } c = \text{Value of FRN}$$

But an FRN trades at par at $t = 0$, and so the swap rate c must be chosen so that

$$\text{Value of bond with coupon } c = \text{Par}$$

Thus the swap rate (for a given maturity) is none other than the corresponding *par yield.*

$$\text{Swap Rate} = \text{Par Yield}$$

Example 6.2.5 3 months ago, corporation A entered the pay–fixed leg of a 2–year swap on a notional of \$100 million, with par swap rate $c\%$, in return for LIBOR, with payments made semiannually. At that stage, LIBOR was $L\%$. To calculate the value of the swap today, use the decomposition of a pay–fixed swap as an FRN minus a fixed coupon bond. The value of the fixed coupon bond today is

$$\text{Bond } = 100 \text{ million} \left[\frac{c}{2}B(0,1/4) + \frac{c}{2}B(0,3/4) + \frac{c}{2}B(0,5/4) + (1+\frac{c}{2})B(0,7/4)\right]$$

In 3 months' time, the FRN will pay a coupon of 100 million $\times\, \frac{L}{2}$, immediately after which it will trade at par. Hence the value of the FRN today is

$$\text{FRN } = 100 \text{ million} \left[1 + \frac{L}{2}\right] B(0,1/4)$$

The value of the swap can now be calculated as FRN – Bond.

□

As for FRA's, we can consider an altered notion of duration for swaps, based on the notional amount rather than the value of the swap (which is initially zero): A fixed–for–floating swap is equivalent to a portfolio which is long a fixed coupon bond and short an FRN. When interest rates change, the value of the FRN changes hardly at all, since FRN's have very small duration. Ignoring the FRN, we see that the duration of a fixed–for–floating swap must be close to that of the fixed coupon bond.

Swaps and Credit Risk

So far, we've ignored the credit risk inherent in swaps, and we briefly consider this now. We've already seen that receive–fixed swap can be regarded as a portfolio of short FRA's, but each with the same fixed rate. Unlike normal FRA's, the FRA's that constitute a swap will have value. Some will have positive value, and some will have negative value. Suppose, for example, that the term structure of interest rates is increasing, and that we enter a pay–fixed swap on a notional of \$1.00, with payments made at times $T_1, \ldots, T_n$. Thus our swap is a portfolio of n FRA's, each with the same fixed rate $x\%$. The swap rate $x\%$ is chosen so that the portfolio of FRA's has zero value. Suppose that the actual forward rate corresponding the k^{th} FRA, i.e. the period $[T_{k-1}, T_k]$, is F_k. The value of the k^{th} FRA that makes up the swap is therefore $(F_k - x) \times (T_k - T_{k-1})B(0, T_k)$. If the swap is to have zero value, it is therefore clear that x must be some kind of "average" of the F_k. Because the term structure is increasing, this "average" is greater than the short term rates, but smaller than the long term rates. Thus the first few FRA's have negative value to pay–fixed: Pay–fixed is paying too much. The final few FRA's have positive value to pay–fixed. Pay–fixed therefore receives negative payments in the short run, and positive payments in the long run.

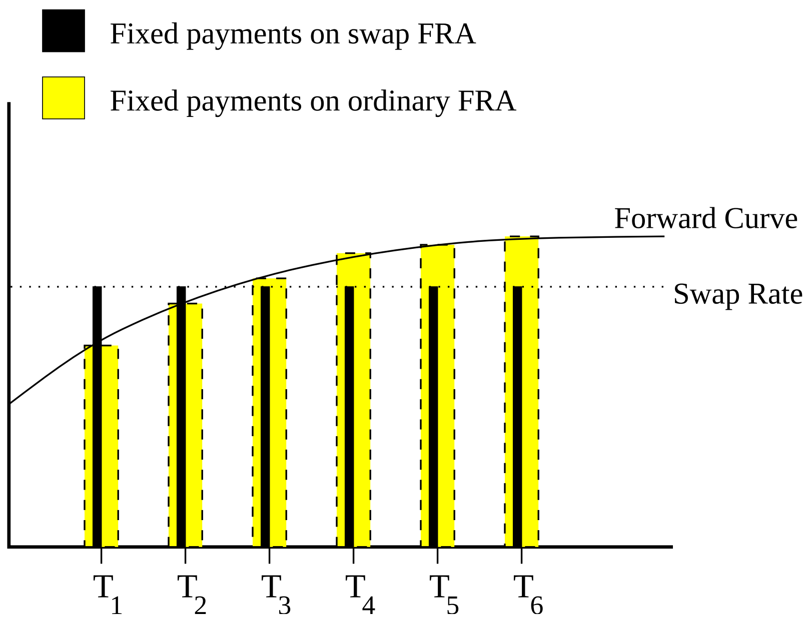

Figure 6.4: Swap rate and forwards rates

If, however, receive–fixed defaults some time during the life of the swap, then pay–fixed has made a number of payments for which it will not be compensated. Thus if the term structure is increasing, pay–fixed is subject to greater credit risk than receive–fixed.

The opposite is true if the term structure is inverted: In that case the swap rate is smaller than short term rates but greater than long term rates.

6.2.5 Other Types of Interest Rate Swap

- In a *basis swap*, two parties interchange floating rate payments, where the floating rates are determined by different indexes. For example, a bank might receive the prime rate on loans that it has made, these loans themselves being funded by interbank borrowing at LIBOR rates. The bank faces the risk that the spread between the prime rate and LIBOR changes (decreases) — basis risk. This can be hedged with a basis swap, exchanging prime for LIBOR (plus a few bp, say).
  Another example is a *constant yield swap* where one rate is linked to yields on long maturity bonds, and another is linked to yields on medium term bonds.

- In a *zero coupon swap*, pay–fixed does not pay until maturity, but receives interim floating rate payments. Clearly pay–floating is subject to greater credit risk.

- A *forward swap* is a swap contract that starts not today, but at some time in the future. The contract swap rate is decided today, however. Suppose, for example, that a corporation is committed to issuing a fixed coupon corporate bond in 6 months' time, but that it prefers to have floating rate debt. It can enter into a forward swap, and hedge against the swap rate moving in an unfavourable direction. Thus a forward swap is essentially a forward on a swap.

- With an *amortising swap*, the notional principal amount reduces over the life of the swap. For example, if a company has taken an amortizing fixed rate loan, the interest payments tend to become smaller. Thus in this case, an amortizing swap is a better hedge than a plain vanilla swap.
  Similarly, *accreting swaps* are swaps where the notional principal amount increases over time. In *rollercoaster swaps*, the notional is allowed to increase and decrease over time, in line with the projected future debt of the customer.

- Swap contracts also come with options attached: A *call swaption* is the option to enter the pay–fixed leg of a swap. The strike rate is the contract "swap" rate, which will generally differ from the prevailing swap rate at the swaption's maturity. A call swaption would clearly be rationally exercised only if the strike rate is less than the prevaling swap rate.
  Similarly, a *put swaption* is the option to enter as a floating rate payer. Of course, these options are not free.

- A *deferred rate setting swap* is like a forward swap. However, the contract swap rate is not decided today, but at any time up to some given future date. Thus the fixed rate payer can defer the setting of the fixed rate.

- A *callable swap* gives the pay–fixed side the right to exit the swap contract, which it might want to do if interest rates drop considerably. This option is paid for by an increase in the fixed rate, i.e. the pay–fixed side of a callable swap pays a higher rate than it would have if this was a plain vanilla swap.
  Puttable swaps give the pay–floating side the right to terminate the swap contract.

- In a *rate–capped swap*, the rate that the pay–floating side may have to be is capped. This is an option belonging to pay–floating, who must pay a premium to pay–fixed at the initiation of the swap contract.

6.3 Interest Rate Futures

Swaps and FRA's are OTC instruments; futures are exchange traded, and interest rate futures are the most commonly traded futures of all. In September 2002, there were 24.4 million futures contracts outstanding (worldwide) of which 16.8 million were interest rate futures[2]. Interest rate futures have been around since 1975, when CBOT introduced futures on Ginny Maes[3] We will briefly cover the two most important types of interest rate futures, namely Eurodollar futures and Treasury bond futures.

6.3.1 T–bond Futures

Recall that futures contracts are agreements to make or take delivery of a fixed amount of a specified grade of asset at some agreed future date for an agreed price, i.e. futures contracts are standardized. Thus a T–bond future is an agreement to make or take delivery of a Treasury bond. To standardize the contract, each T–bond future calls for the *physical delivery* of a

[2]Source: BIS Quarterly Review, December 2002

[3]"Ginny Mae" comes from GNMA, the *Government National Mortgage Association.* Ginny Maes are USA mortgage–backed long term bonds.

Treasury bond with face value of \$100 000 and at least 15 years to maturity. It must be non–callable within 15 years as well. However, different bonds may have different coupon rates. The *standard* is currently a 6% coupon, although this may vary: CBOT changed the notional coupon from 8% to 6% in March 2000. To ensure the interchangeability of bonds, each T–bond is assigned a *conversion factor* which converts the quoted futures price to the price that is actually paid by the party who is long the future. This is important, because different bonds trade at different prices. High coupon bonds trade at a premium, whereas low coupon bonds trade at discount. As we shall see, the conversion factor for a bond with a particular coupon is determined by pretending that the yield curve is flat at 6%. Accrued interest must also be taken into account.

Like T–bonds themselves, T–bond futures are quoted in 32^{nds} of dollars, i.e. a futures price of 92-25 means a price of $\$92\frac{25}{32}$ per \$100 of face value.

Let's have a look at the mechanics of a T–bond futures contract. The party who is short the future must deliver, at any time during the delivery month, a T–bond with maturity at least 15 years on the first day of the delivery month (and non–callable within 15 years as well). There are many such bonds, and the short side gets to choose which bond will be delivered. The long side will then pay a cash amount which is determined by the quoted futures price, the conversion factor of the delivered T–bond and the accrued interest since the last coupon of the delivered T–bond:

$$\text{Cost} = \text{Quoted Price} \times \text{Conversion Factor} + \text{Accrued Interest}$$

Example 6.3.1 Suppose that on expiry of the futures contract, the quoted futures price is 92-25. The party short the future decides to deliver a bond with conversion factor of 1.2406. The accrued interest on the delivery date is \$3.1157 per \$100 of face value. The cash paid by the party who is long is

$$\$92\frac{25}{32} \times 1.2406 + 3.1157 = 118.22012$$

per \$100 of notional. Since the face value of a T–bond is \$100 000, the actual cash paid is \$118 220.12

$\square$

The conversion factor for a bond is calculated as follows:

- Calculate the theoretical value of the bond on the *first day of the delivery month* of the futures contract, subject to the following:
  - Assume that the term structure is *flat* at 6% (with semiannual compounding).
  - Assume that the bond maturity and the coupon dates are *rounded down* to the nearest 3 months.
  - If (after rounding) the bond maturity is a multiple of 6 months, assume that the first coupon is paid in 6 months time, and then at 6 month intervals until the rounded maturity.
  - If (after rounding) the bond maturity is not a multiple of 6 months (in which case it is a multiple of 3 months), assume that the first coupon is paid in three months time. Thus it is assumed that the bond is exactly halfway in between coupons. In that case, find the price of the bond in 3 months time (including the coupon), discount it back 3 months (at 6% p.a. semiannually) and subtract half the coupon.

This procedure makes it possible for CBOT to publish tables of conversion factors, rather than letting investors figure it out for themselves.

Example 6.3.2 Consider a T–bond future for delivery in March 2003. Suppose that the short side decides to deliver the November 2028 $5\frac{1}{2}\%$ T-bond. In March 2003, the T–bond will have a maturity of 25 years and 8 months. This is rounded down: We thus calculate the bond price as though it matures in 25 years and 6 months. The first coupon is assumed to be paid in 6 months time, so the bond price is

$$\begin{aligned} P &= \sum_{k=1}^{51} \frac{2.75}{1.03^k} + \frac{100}{1.03^{51}} \\ &= \frac{2.75}{1.03} \frac{1 - \frac{1}{1.03^{51}}}{1 - \frac{1}{1.03}} + \frac{100}{1.03^{51}} \\ &= 93.51 \end{aligned}$$

The conversion factor is therefore 0.9351. This is less than 1, because the bond's coupon is less than 6%.

□

Example 6.3.3 Consider a T–bond future for delivery in July 2003. Suppose that the short side decides to deliver the December 2020 8% coupon bond. In July 2003, the T–bond will have a maturity of 17 years and 5 months. Rounding down, we regard the maturity as being in 17 years and 3 months. Thus there are 35 coupons left. In 3 month's time, the bond price will be

$$\begin{aligned} P &= \sum_{k=0}^{34} \frac{4}{1.03^k} + \frac{100}{1.03^{34}} \\ &= 4 \frac{1 - \frac{1}{1.03^{35}}}{1 - \frac{1}{1.03}} + \frac{100}{1.03^{34}} \\ &= 125.1318 \end{aligned}$$

This must be discounted 3 months back. Now the discount factor is $\frac{1}{1.03^{\frac{1}{2}}} = 0.95329$, so the discounted bond price is \$123.30. Since we are halfway between coupon dates, we are entitled to accrued interest amounting to \$2.00, so the bond price is \$121.30. It follows that the conversion factor is 1.2130.

□

Because there are many different kinds of bond that can be delivered in a T-bond futures contract, the party which is short the future can choose the bond which is cheapest to deliver. The short side receives

$$\text{Futures Price} \times \text{Conversion Factor} + \text{Accrued Interest}$$

The cost of the bond delivered is

$$\text{Quoted Bond Price} + \text{Accrued Interest}$$

Thus the gain made by the short side is

$$\text{Gain} = \text{Futures Price} \times \text{Conversion Factor} - \text{Quoted Bond Price}$$

The bond for which this gain is a maximum is called *the cheapest–to–deliver* bond.

Example 6.3.4 Suppose that quoted bond prices and their conversion factors are as follows:

Bond	Price	Conversion Factor
A	162-20	1.3987
B	138-31	1.2870
C	131-2	1.2730

Also assume that the quoted T–bond futures price is 114-26. Which bond is cheapest to deliver?

The short side's gain on each of the bonds is easily calculated. For bond A, the gain is -2.037; for bond B, 8.795; for bond C, 15.094. Thus bond C is cheapest to deliver.

□

Whereas it is possible to determine a theoretical arbitrage–free futures price for equity and commodity futures, this is more difficult in the case of T–bond futures, because the short party has some choices as to when to deliver (at any time during the delivery month) and what to deliver (any T–bond with maturity $\geq$ 15 years). However, if both the delivery date and the T–bond are known, a theoretical futures price *of that bond* can be calculated in the usual way:

$$F_0 = \text{FV}(S_0 - I)$$

where S_0 is the spot price, and I is the present value of all cash flows (coupons) during the life of the future. However, this is not the whole story; the fact that quoted prices are not spot prices, and that there are conversion factors must also be taken into account. Given today's quoted bond price, one has to find the spot price by adding the accrued interest. One then must subtract the present values of any coupons paid during the life of the option (which requires knowledge of the yield curve). Taken the future value of this would give the "actual" futures price of the bond, so accrued interest must be subtracted to find the *quoted futures price of the bond*. This must then be divided by the conversion factor of the bond to give an estimate of the quoted T-bond futures price.

In addition to Treasury bond futures, CBOT also trades futures on 10–year, 5–year and 2–year Treasury notes.

6.3.2 Eurodollar Futures

A Eurodollar is a US dollar–denominated deposit held by a bank outside the USA. The rates at which such a bank is willing to lend dollars is just LIBOR, and Eurodollar futures are essentially bets on LIBOR movements. Traded on the CME, they are similar to FRA's on 3–month forward rates, and have maturities up to ten years.

Each futures contract is written on a 3–month Eurodollar deposit of $1 000 000, but all contracts are *cash settled*. Futures prices are quoted as follows:

$$\text{Quoted Futures Price} = 100 - \text{Futures Rate } (\%)$$

For example, if the quoted futures price is 95.20, then the *futures discount rate* is 4.80%. The price of a contract is therefore

$$1\,000\,000\left[1-\frac{0.0480}{4}\right]=\$988\,000$$

This is the same as

$$10\,000\,[100-0.25(100-\text{Futures Price})]$$

Note that if the quoted futures price goes up by 0.01 (i.e. the futures rate goes down by 0.01% = 1bp), the contract price increases by $25.

As with all futures, Eurodollar futures are marked to market, so that multiples of $25 exchange hands every day. The final marking–to–market sets the quoted futures price to 100 − LIBOR, where LIBOR is the 90–day LIBOR spot rate on the delivery date.

Example 6.3.5 We give here a rough example of a hedging strategy. Suppose that a bond trader wants to hedge a $100 million portfolio of 1–year Treasury bills. Treasury bills are zero coupon bonds, so the duration of the portfolio is 1 year. Hence if interest rates go up by 1 bp, the portfolio loses roughly $10 000. To match this with a position in Eurodollar futures, recall that if interest rates go up by 1bp, the Eurodollar contract price decreases by $25. Thus the bond trader should take a short position of approximately $\frac{10\,000}{25}=400$ Eurodollar futures.

□

6.4 Interest Rate Options

In this section, we briefly discuss some popular interest rate options. The pricing of such options requires heavy mathematical machinery, and is still a very active area of research.

A *bond call option* gives its holder the right to buy a particular bond on a particular date (or set of dates) for a particular strike price. Similarly, a bond put option confers on its holder the right to sell. Bond options can be European or American and are generally traded OTC. Note that some bonds already have option features built into them, e.g. the callable and puttable bonds discussed previously.

In South Africa, bonds are quoted in yields, not price. The same applies to the strike of a bond option, i.e. instead of the right to buy the R150 for R900 000, the holder of an option may have the right to buy the R150 at a yield of 16.87%. This makes no real difference for European options, because one can always calculate what price corresponds to what yield on the maturity of the option. It does, however, make a difference when we consider American bond options: Since the price of a bond will vary from day to day even if the yield is constant, the strike price of an American bond option struck on yield is not constant.

An *interest rate cap* protects its holder against rising interest rates. Consider, for example, a corporation that has to pay LIBOR on a 5–year loan of $1 million, with payments made quarterly. Suppose that LIBOR is 12% today, and that management believe that the corporation might face bankruptcy if LIBOR goes above 15%. Management can protect the corporation by buying an interest rate cap on a notional of $1 million, with *cap rate* 15% and *tenor* 3 months. If, for example, in a year's time LIBOR is 14%, then the seller of the cap need not pay the corporation anything. Suppose, however, that in two years' time LIBOR

has risen to 16%. The corporation is then responsible for a quarterly interest payment of \$40 000. But at that time they will also receive

$$\begin{aligned}\text{Payment on Cap} &= \text{Notional} \times (\text{LIBOR} - \text{Cap Rate}) \times \text{Tenor} \\ &= \$1\,000\,000 \times 0.01 \times 0.25 \\ &= \$2\,500\end{aligned}$$

from the seller of the interest rate cap (who might not be the same as the party to whom interest payments are due). The total cash flow in 2 years' time is therefore \$37 500, i.e. 15% of \$1 000 000 over a quarter. Thus any interest above the 15%–level is owed by the seller of the cap.

An interest rate cap is essentially a portfolio of call options on a floating reference rate. Each such call option is called a *caplet*. Consider a cap on a notional amount of F with a cap rate of c. Suppose that the floating rate is reset n times during the life of the cap, at times $t_1, \ldots, t_n$, and let $t_{n+1} = T$, the expiry of the cap. Then the cap consists of n caplets. The k^{th} caplet protects against rising prices over the k^{th} period, i.e. $[t_k, t_{k+1}]$. If LIBOR(k) is the floating rate over the period $[t_k, t_{k+1}]$, then at time t_{k+1}, the seller of the cap must pay the holder

$$\text{Payment on Caplet} = F\Delta_k \left[\text{LIBOR}(k) - c\right]^+$$

where $\Delta_k = t_{k+1} - t_k$ is the length of the time period. Note that this is an *in–arrears cap*, i.e. the payment is made at the end of the corresponding period, when interest payments are due. (An in–advance cap would, at time t_k, pay the above amount, discounted to time t_k.)

A cap can also be seen as a portfolio of put options on zero coupon bonds: A caplet payment of $F\Delta_k \left[\text{LIBOR}(k) - c\right]^+$ at time t_{k+1} is equivalent to a payment of

$$\frac{F\Delta_k}{1 + \text{LIBOR}(k)\Delta_k} \left[\text{LIBOR}(k) - c\right]^+$$

at time t_k. However,

$$\frac{F\Delta_k}{1 + \text{LIBOR}(k)\Delta_k} \left[\text{LIBOR}(k) - c\right]^+ = \left[F - \frac{F(1 + c\Delta_k)}{1 + \text{LIBOR}(k)\Delta_k}\right]^+$$

It can now be seen that the caplet is equivalent to a put option with strike F and expiry t_k on a zero coupon bond with face value $F(1 + c\Delta_k)$ and maturity t_{k+1}[4].

An *interest rate floor* protects its holder against dropping interest rates. It can be regarded as a portfolio of put options on a floating reference rate. Each such option is called a *floorlet*, and has a payoff of

$$\text{Payment on Floorlet} = F\Delta_k \left[c - \text{LIBOR}(k)\right]^+$$

at time t_{k+1} (in–arrears). A floor can also be regarded as a portfolo of call options of call options on zero coupon bonds.

Note that there is a put–call parity relation between caps and floors with the same notional, cap/floor rate and tenor: Consider a portfolio which is long a cap and a receive–fixed vanilla swap (with the same notional and tenor, and with fixed rate equal to the cap rate c). Now consider the following possibilities:

[4] This is because the time–t_k value of a zero coupon bond with face value $F(1 + c\Delta_k)$ and maturity t_{k+1} is $\frac{F(1+c\Delta_k}{1+\text{LIBOR}(k)\Delta_k)}$.

- If LIBOR $> c$, then the cap pays LIBOR $- c$ and the swap pays $c -$ LIBOR. Hence the portfolio pays zero.
- If LIBOR $\leq c$, then the cap pays nothing, and the swap pays $c -$ LIBOR. Hence the portfolio pays $c -$ LIBOR.

Thus in both cases, the portfolio pays the same as a floor (with the same notional, tenor, and floor rate c). We therefore have

$$\text{Floor} = \text{Cap} + \text{Receive–Fixed Swap}$$

An *interest rate collar* ensures that interest payments remain within certain bounds, c_{up} and c_{down}. If $c_{\text{down}} \leq \text{LIBOR} \leq c_{\text{up}}$, then the holder of the collar pays and receives nothing. If LIBOR $\geq c_{\text{up}}$, then the holder of the cap *receives* LIBOR $- c_{\text{up}}$. If LIBOR $\leq c_{\text{down}}$, the holder of the collar *pays* $c_{\text{down}} -$ LIBOR. Thus a collar is a long cap and a short floor:

$$(\text{Collar with bounds } c_{\text{down}} \leq c_{\text{up}}) = (\text{Cap with rate } c_{\text{up}}) - (\text{Floor with rate } c_{\text{down}})$$

The rates $c_{\text{up}}, c_{\text{down}}$ are often chosen so that the initial value of the collar is zero. Note that a swap is just a special case of this: A pay–fixed swap with fixed rate c is equivalent to a collar with $c_{\text{up}} = c_{\text{down}} = c$.

These are just some of the most popular interest rate options. Of course, there are many more. We have already mentioned swaptions (the option to enter a swap). This is a 2nd order derivative, i.e. a derivative on a derivative. Another example of a 2nd order interest rate derivative is a *caption*, which is an option to enter into a cap at some future date. There are also options on bond futures and Eurodollar futures.

6.5 Duration–Based Hedging Strategies

In Section 3.3, we discussed futures–based hedging and determined the optimal hedge ratio. We will now look at hedging in the setting of interest rate risk.

Consider a portfolio of debt instruments with value P and modified duration D_P^*, so that approximately

$$\Delta P = -D_P^* P \Delta y$$

where Δy is a small parallel shift in the yield curve.

Also consider a position in futures F, with modified duration D_F^*, where F is the futures price, and D_F^* is defined so that

$$\Delta F = -D_F^* F \Delta y$$

We now want to add futures to our portfolio P to form a new portfolio Π which has a target modified duration D_Π^*. For example, we could aim for $D_\Pi^* = 0$ if we wanted to remove first order interest rate risk.

Suppose that a risk manager adds n futures to the portfolio P, so that $\Pi = P + nF$. Thus

$$\Delta\Pi = \Delta P + n\Delta F \qquad \text{i.e.}$$
$$-D_\Pi^* \Pi \Delta y = -(D_P^* P + n D_F^* F)\Delta y$$

Thus

$$n = \frac{D_\Pi^* \Pi - D_P^* P}{D_F^* F}$$

is the number of futures that must be used to attain a target duration of D^*_Π.

The quantity $n = -\frac{D^*_P P}{D^*_F F}$ sets the duration of Π to zero. It is called the *duration–based hedge ratio.*

Example 6.5.1 Suppose that a bond portfolio manager must hedge a \$100 million portfolio for 3 months. The duration of the portfolio is 7.3 years, and the target duration is 0. The manager wants to use T–bond futures to effect the hedge. Currently, the futures price is $94\frac{8}{32}$ (notional \$100 000). The future's duration can be estimated by that of the cheapest–to–deliver, which is 8.4 years.

From the above, the number of futures added to the portfolio must be

$$n = -\frac{D^*_P P}{D^*_F F} = -\frac{7.3 \times 100 \text{ million}}{8.7 \times 94\,250} = -922.07$$

i.e. the manager should short approximately 922 futures contracts.

□

Example 6.5.2 Here is an example of a rough hedge: Today is July 12. A corporation has undertaken to issue commercial paper with a maturity of 6 months on October 20, which will provide an expected income of \$5.3 million. The corporate treasurer is worried that rates will increase, because this will lead to less income from the issue. If rates increase by 1 basis point, the value of the issue of commercial paper will decrease by 5.3 million $\times$ 0.5 $\times$ 0.01% = \$265, since the duration of a zero coupon bond equals its time to maturity. (Here we are ignoring the difference between duration and dollar duration).

To hedge this risk, the treasurer decides to use Eurodollar futures. If rates increase by 1 basis point, the futures price of a Eurodollar futures contract decreases by \$25. Thus the duration of a Eurodollar futures contract is 3 months. In order to match a loss of \$265 with a gain on futures, the treasurer should short approximately $\frac{265}{25} = 10.6$ contracts, i.e. 11 December Eurodollar futures.

□

Chapter 7

The Black–Scholes Model

In this chapter we derive the Black–Scholes equations for European options. The Black–Scholes price is not model independent, i.e. it depends on the model we chose for stock prices. Accordingly, the first section of this chapter is concerned with developing a model of stock price behaviour. In the second section, we develop the machinery of stochastic calculus in an intuitive and non–rigorous manner. It should be pointed out that the motivation for Itô's formula provided in this section is pretty flimsy, and we do not claim to give a mathematically accurate account.

Having Itô's formula at our disposal, we then derive the Black–Scholes PDE, again in an intuitive non–rigorous manner. Instead of solving the PDE directly — we will do that later — we note that the PDE has a surprising property: The drift of the underlying asset does not occur in the PDE. This allows us to use the machinery of *risk–neutral valuation* to derive the Black–Scholes option prices.

7.1 Modelling Stock Prices

Any model of stock price behaviour must be *stochastic*, i.e. it must incorporate the random nature of price behaviour. The simplest such models are *random walks*: Let $X_t, t = 1, 2, \ldots$ be a family of distributed random variables, and let S_0 be the stock price at $t = 0$. We might (naively) attempt to model the stock price process by

$$S_t = S_{t-1} + X_t \quad \text{i.e.} \quad S_t = S_0 + \sum_{u=1}^{t} X_u$$

The intuition behind this is that the price at time t equals the price at time $t-1$ plus a "random shock", modelled by X_t.

We should also assume that these shocks are *independent*. Why? If we could predict today that the stock price is going to go up *tomorrow*, this makes the stock more attractive today. Thus more people would buy it today, forcing the stock price up *today*, until it reaches the level predicted. Thus any change in the stock price must essentially be unpredictable. This is just a version of *Efficient Markets Hypothesis*, which, loosely, asserts that all available information about a corporation is instantly reflected in its stock price. Thus future changes in price are not dependent on past changes in price.

There are several reasons why a random walk model of stock prices is inadequate, but an obvious one is that it doesn't take into account scale. for stock prices, we expect the change

in price to be *proportional* to the current price. To see this, consider two companies in two parallel universes, A and B. The universes and the companies are identical, except for one thing. In universe A, the company has issued 100 shares, each trading at $100. In universe B, the company has undertaken a 2–for–1 stock split, so that it has issued 200 shares, each trading at $50. Both companies are otherwise identical, e.g. they are both worth $10 000. One day an earthquake cause massive damage, and both companies lose half their value. The shares in universe A now trade at $50, whereas those in universe B trade at $25. Thus the share price has not dropped by the same amount in both universes: Each share has lost the same *proportion* of its value.

Simply put, if investors require a return of 14%, then they require that return irrespective of whether the share price is $50 or $100.

The shares of A, B change by the same *factor*, i.e. they have exactly the same change in returns (but not the same absolute change in price). This is reflected in, e.g., the binomial model, where shares can go up by a *factor* of u or down by a *factor* of $\frac{1}{u}$. But a multiplicative change in the stock price amounts to an additive change in the logarithm of the stock price:

$$S_{t+\Delta t} = u^{\pm 1} S_t \qquad \text{implies} \qquad \ln S_{t+\Delta t} = \ln S_t \pm \ln u$$

i.e. if we define the returns process R_t by $S_t = S_0 e^{R_t}$ (i.e. $R_t := \ln \frac{S_t}{S_0}$), and define $\delta := \ln u$, we have i.e.

$$R_{t+\Delta t} = R_t \pm \delta$$

A better random model of stock prices is therefore one in which the *returns* process R_t follows a random walk.

7.1.1 Modelling Returns in Continuous–Time

We now seek a continuous-time version of the random walk — a stochastic process that is changing because of random shocks at every instant in time. Consider a time interval $[0, T]$ and let N be a (large) integer. Define $\Delta t := \frac{T}{N}$. Let $X_n, n = 1, 2, 3, \ldots$ be independent Bernoulli random variables with

$$\mathbb{P}(X_n = \Delta x) = p \qquad \text{and} \qquad \mathbb{P}(X_n = -\Delta x) = 1 - p =: q$$

where $\Delta x > 0$. For $t = 0, \Delta t, 2\Delta t, \ldots, N\Delta t = T$, let $R_t := \sum_{i=1}^{n} X_n$, where $t = n\Delta t$. Thus R_t is a random walk, and

$$R_{t+\Delta t} = R_t \pm \Delta x$$

Some simple calculations yield

$$\mathbb{E}[R_t] = n(p-q)\Delta x = (p-q)\frac{\Delta x}{\Delta t}t \qquad \mathrm{Var}(R_t) = n(\Delta x^2 - (p-q)^2\Delta x^2) = 4pq\frac{\Delta x^2}{\Delta t}t$$

Now suppose we can observe the process R_t and want $\mathbb{E}[R_t] = \mu t$ and $\mathrm{Var}(R_t) = \sigma^2 t$, where μ, σ are constants, and $\sigma > 0$. (We want $\sigma^2 > 0$, otherwise $\mathrm{Var}(R_t) = 0$, in which case R_t would be non–random.)

In the continuous limit, i.e. as $N \to \infty$ and $\Delta t \to 0$, we must have

$$(p-q)\frac{\Delta x}{\Delta t} \to \mu \qquad 4pq\frac{\Delta x^2}{\Delta t} \to \sigma^2$$

The first equation yield $\Delta x \approx \frac{\mu\Delta t}{p-q}$ when Δt is small. Substituting into the second equation, we see that

$$\frac{4pq}{(p-q)^2}\Delta t \approx \frac{\sigma^2}{\mu^2}$$

when Δt is small. Now since, $\Delta t \to 0$, we must have $\frac{4pq}{(p-q)^2} \to \infty$, for otherwise the product $\frac{4pq}{(p-q)^2}\Delta t$ would tend to 0, not $\frac{\sigma^2}{\mu^2}$. It is therefore necessary that $p - q \to 0$, and thus p, q must both tend to $\frac{1}{2}$ as $\Delta t \to 0$. From the fact that $4pq\frac{\Delta x^2}{\Delta t} \to \sigma^2$, we then see that we must have

$$\Delta x \approx \sigma\sqrt{\Delta t}$$

for small Δt.

We had $\Delta x \approx \frac{\mu\Delta t}{p-q}$ for small Δt, and thus $p - q \approx \frac{\mu}{\sigma}\sqrt{\Delta t}$. Since $p + q = 1$, we must have

$$p = \tfrac{1}{2}(1 + \tfrac{\mu}{\sigma}\sqrt{\Delta t}) \qquad = \tfrac{1}{2}(1 - \tfrac{\mu}{\sigma}\sqrt{\Delta t})$$

As a check, note that

$$\mathbb{E}[R_t] = (p-q)\frac{\Delta x}{\Delta t}t = \frac{\mu}{\sigma}\sqrt{\Delta t}\frac{\sigma\sqrt{\Delta t}}{\Delta t}t = \mu t$$

and

$$\mathrm{Var}(R_t) = 4pq\frac{\Delta x^2}{\Delta t}t = (1 - \tfrac{\mu^2}{\sigma^2}\Delta t)\frac{\sigma^2\Delta t}{\Delta t}t = \sigma^2 t - \mu^2 t\Delta t \to \sigma^2 t$$

as should be the case.

We now have an idea of how to create a a continuous–time stochastic process R_t as the $(\Delta t \to 0)$–limit of a random walk. But the limit process has some peculiar features. For example

$$\Delta R_t \approx \pm\sigma\sqrt{\Delta t} \quad \text{is of the order of } \sqrt{\Delta t}$$

If $f(t)$ is a differentiable function, then

$$\Delta f(t) \approx f'(t)\Delta t \quad \text{is of the order of } \Delta t$$

. Now when Δt is small, we see that $\sqrt{\Delta t}$ is much larger than Δt (Take, e.g. $\Delta t = 10^{-2n}$ and note that $\sqrt{\Delta t} = 10^{-n} = 10^n\Delta t$.) It follows that R_t cannot be differentiable as a function of t.

The probabilist will immediately want to know the distribution of R_t. Let $u(t, x)$ be the density of the random variable R_t, i.e.

$$u(t, x)\Delta x \approx \mathbb{P}(R_t \in [x, x + \Delta x])$$

At time $t + \Delta t$ the random walk can reach the point x in two ways: It can move right from the point $x - \Delta x$ at time t, with probability p, or it can move left from the point $x + \Delta x$, with probability q. Thus

$$u(t + \Delta t, x) = pu(t, x - \Delta x) + qu(t, x + \delta x)$$

Now we Taylor expand up to order Δt. Firstly

$$u(t + \Delta t, x) \approx u(t, x) + u_t(t, x)\Delta t + o(\Delta t)$$

Next,

$$u(t, x \pm \Delta x) = u(t,x) \pm u_x(t,x)\Delta x + \tfrac{1}{2}u_{xx}(t,x)\Delta x^2 + o(\Delta x^2)$$

Here, we have taken a second–order Taylor expansion, because Δx is of the order $\sqrt{\Delta t}$, and Δx^2 of the order Δt. Putting these together, we obtain (at the point (t,x)):

$$u + u_t\Delta t = (p+q)u + (-p+q)u_x\Delta x + \tfrac{1}{2}(p+q)u_{xx}\Delta x^2$$

However, we know that $p = \frac{1}{2}(1 + \frac{\mu}{\sigma}\sqrt{t})$ and that $\Delta x \approx \sigma\sqrt{\Delta t}$ and $p, q \to \frac{1}{2}$. Hence

$$u_t\Delta t = -(\tfrac{\mu}{\sigma}\sqrt{\Delta t})u_x(\sigma\sqrt{\Delta t}) + \tfrac{1}{2}u_{xx}\sigma^2\Delta t$$

which yields the following partial differential equation for the density of R_t.

$$u_t = -\mu u_x + \tfrac{1}{2}\sigma^2 u_{xx}$$

However, the PDE is not sufficient to determine the density u: It has many solutions. We seek a solution which has the following properties:

- For each $t \geq 0$, we have $\int_{-\infty}^{\infty} u(t,x)\,dx = 1$, because $u(t,x)$ is a density, and
- $u(0,x)$ is rather odd: We have $R_0 = 0$, and so

  $$f(0) = \mathbb{E}[f(R_0)] = \int_{-\infty}^{\infty} f(x)u(0,x)\,dx$$

  i.e. $u(0,x)$ is a "function" with the property that $\int_{-\infty}^{\infty} f(x)\,dx = f(0)$ for every function f. The "function" with this property is called the *Dirac delta* δ_0. It is not a function at all (but the simplest example of a so–called *generalized function* or *distribution* (in the sense of Schwartz).) Nevertheless, we can get some intuition as to how u ought to behave. We see that for t close to 0, the density $u(t,x)$ must be very small for $x \neq 0$, because R_t must be close to x when t is near zero. Yet the area under the curve is 1, i.e. $u(t,x)$ must be extremely peaked at around $x = 0$ and then rapidly drop off. We may thus think off $u(0,x) = \delta_0$ as a "function" which has

  $$\delta_0(x) = 0 \text{ when } x \neq 0 \qquad \delta_0(0) = +\infty \text{ in such a way that } \int_{-\infty}^{\infty} \delta_0(x)\,dx = 1$$

Oddly enough, we can find such a function. The PDE for the density, derived by Einstein in 1905, is a version of the heat equation, derived by Fourier, which governs heat transfer. So this PDE was not new: It had been intensively studied by physicists, with $u(t,x)$ playing the role of the temperature at time t at a point x in an infinitely long rod. The *fundamental solution* or *Green's function* of such a PDE was well–known

$$u(t,x) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}$$

We will give a derivation of this result later on, but you can verify by direct differentiation that this function does, in fact, satisfy the PDE. You will also immediately recognize it as the density of an $N(\mu t, \sigma^2 t)$–random variable. Furthermore, for t near 0, such a random variable

has very small standard deviation, and thus the density is extremely peaked around 0, just as we require.

It follows, therefore, that the density of $_t$ is $N(\mu t, \sigma^2 t)$. Of course, the Central Limit Theorem states that, subject to a moment condition, large sums of i.i.d. are roughly normally distributed, so we are not surprised. But here, we have in essence given a proof of the Central Limit Theorem by PDE methods, at least for random walks of the type described.

When we take $\mu = 0$ and $\sigma = 1$, we obtain one of the basic building blocks of financial modelling:

Definition 7.1.1 Standard *Brownian motion* is a continuous–time stochastic process $B_t, t \geq 0$ with the following properties:

(1) Each change

$$B_t - B_s = (B_{s+h} - B_s) + (B_{s+2h} - B_{s+h}) + \cdots + (B_t - B_{t-h})$$

is *normally distributed* with mean 0 and variance $t - s$.

(2) Each change $B_t - B_s$ is **independent** of all the previous values $B_u, u \leq s$.

(3) Each sample path $B_t, t \geq 0$ is (a.s.) **continuous**, and has $B_0 = 0$.

Now put

$$R_t = \mu t + \sigma B_t$$

It then follows easily that

$$R_t \sim N(\mu t, \sigma^2 t)$$

i.e. the standard Brownian motion can also be used to to model returns processes where $\mu \neq 0$ and $\sigma \neq 1$. The process R_t is called an *arithmetic Brownian motion* with drift rate μ and variance rate σ^2. We will also refer to σ as the *volatility*.

7.1.2 Modelling Share Prices in Continuous Time

We have obtained a model for share prices:

$$S_t = e^{R_t} = e^{\mu t + \sigma B_t}$$

We shall soon see that this translates to a *stochastic differential equation*

$$dS_t = \alpha S_t \, dt + \sigma S_t \, dB_t \qquad \alpha := \mu + \tfrac{1}{2}\sigma^2$$

i.e. the proportional change in share price $\frac{dS_t}{S_t}$ can be decomposed into two terms, $\alpha \; dt$ and $\sigma \; dB_t$. Such a process is called a *geometric Brownian motion* (GBM) with *drift* α and volatility σ., The drift is the (proportional) rate at which the share price increases in the absence of risk. The differential dB_t models the randomness (risk), and the volatility models how sensitive the share price is to these random events. The greater α, the faster the share price increases in the absence of risk. The greater σ, the more violently the share price reacts to random events. Note that dB_t can be negative (unlike dt), allowing for decreases in share price. Also note that $|dB_t| \approx \sqrt{dt} >> dt$, so that over short periods the change in share price is dominated by random events. Many of these random events cancel out however, so that in the long run the drift term is dominant.

Now consider a market with a share S_t whose price follows a GBM $dS_t = \alpha S\, dt + \sigma S\, dB_t$. Let the risk–free interest rate be r, i.e. the risk–free bank account A_t satisfies the DE

$$dA_t = rA_t\, dt$$

A_t is the *riskless asset*. It has drift r and zero volatility.

A portfolio is a two–dimensional process (θ_t^0, θ_t^1), where θ_t^1 is the number of shares owned at time t, θ_t^0 is the amount of money in the bank account at time t, discounted to time 0. Given such a dynamic portfolio $\theta_t = (\theta_t^0, \theta_t^1)$, the *value process* $V_t(\theta)$ satisfies

$$\begin{aligned} dV_t &= \theta_t^0\ dA_t + \theta_t^1\ dS_t \\ &= (r\theta_t^0 A_t + \mu\theta_t^1 S_t)\ dt + \theta^1 \sigma S_t\ dB_t \end{aligned}$$

The value of the portfolio at time T is therefore

$$\begin{aligned} V_T(\theta) = V_0(\theta) &+ \int_0^T [r\theta_t^0 A_t + \mu\theta_t^1 S_t]\ dt \\ &+ \int_0^T \theta^1 \sigma S_t\ dB_t \end{aligned}$$

We now see that we need to be able to evaluate integrals of the form

$$\int_0^T f(t)\ dB_t$$

This is an example of a *stochastic integral*. The obvious method would be to regard the above as a Riemann–Stieltjes (or Lebesgue–Stieltjes) integral. However, it can be shown that this approach will not work. Nevertheless, it is possible to define the stochastic integrals, and there is even a very simple rule which allows us to manipulate them: Itô's formula. However, the rules of stochastic calculus do differ from those of ordinary calculus. We are, after all, now working with stochastic processes whose paths are nowhere differentiable, whereas ordinary calculus deals only with differentiable functions.

7.2 A Naive Approach to Stochastic Calculus

Let $f(x)$ be a differentiable function on an interval $[a, b]$. Partition this interval:

$$a = x_0 < x_1 < x_2 < x_n = b$$

where $x_{i+1} - x_i = \Delta x$. Then by Taylor series expansion, we get

$$\begin{aligned} f(x_{i+1}) - f(x_i) &= f'(x_i)\Delta x + \frac{1}{2!} f''(x_i)(\Delta x)^2 \\ &+ \frac{1}{3!} f'''(x_i)(\Delta x)^3 + \text{ terms involving } \Delta x^4, \Delta x^5, \ldots \end{aligned}$$

Thus

$$\begin{aligned} f(b) - f(a) &= \sum_{i=0}^{n-1} [f(x_{i+1}) - f(x_i)] \\ &= \sum_{i=0}^{n-1} f'(x_i)\Delta x + \frac{1}{2} \sum_{i=0}^{n-1} f''(x_i)(\Delta x)^2 + \ldots \end{aligned}$$

As $\Delta x \to 0$, we get

$$\begin{aligned}f(b) - f(a) &= \lim_{\Delta x \to 0} \sum_i f'(x_i)\Delta x + \frac{1}{2} \lim_{\Delta x \to 0} \sum_i f''(x_i)(\Delta x)^2 \\ &\quad + \ldots \\ &= \int_a^b f'(x)\ dx + \left[\frac{1}{2}\int_a^b f''(x)\ (dx)^2 + \ldots\right]\end{aligned}$$

In ordinary calculus, only the first term counts (by the Fundamental Theorem of Calculus), and the other terms are zero. This is because the *quadratic variation* of any "ordinary" function is zero, i.e.

$$\lim_{\Delta x \to 0} \sum (\Delta g)^2 = 0$$

for any "ordinary" function g. This is not all that hard to see: We have

$$\begin{aligned}\lim_{\Delta x \to 0} \sum (\Delta g)^2 &= \lim_{\Delta x \to 0} \sum g'(x)^2 \Delta x^2 \\ &= (\lim_{\Delta x \to 0} \Delta x)(\lim_{\Delta x \to 0} \sum g'(x)^2 \Delta x) \\ &= 0 \cdot \int_a^b g'(x)^2\ dx \\ &= 0\end{aligned}$$

(assuming that g is continuously differentiable).

But Brownian motion is different: Consider $\Delta B = B_{t+\Delta t} - B_t$. This is a normally distributed random variable with $\mathbb{E}[\Delta B] = 0$ and variance $\mathrm{var}(\Delta B) = \Delta t$.

Consider next the random variable $(\Delta B)^2$. This has

$$\begin{aligned}\mathbb{E}[(\Delta B)^2] &= \mathrm{var}[\Delta B] = \Delta t \\ \mathrm{var}[(\Delta B)^2] &= \mathbb{E}[(\Delta B)^4] - (\Delta t)^2 = 2(\Delta t)^2 << \Delta t\end{aligned}$$

Thus the variance of $(\Delta B)^2$ is ≈ 0, i.e. though ΔB is a random variable, $(\Delta B)^2$ is a constant.[1] It follows that

$$\lim_{\Delta t \to 0} \sum \mathbb{E}(\Delta B)^2 = \lim_{\Delta t \to 0} \sum \Delta t = T$$

where T is the total elapsed time. Thus the quadratic variation of Brownian motion is non–zero.

Also

$$\lim_{\Delta t \to 0} \sum \mathbb{E}(\Delta B)^4 = 2 \lim_{\Delta t \to 0} \sum (\Delta t)^2 = 0$$

because $g(t) = t$ is an "ordinary" function, with quadratic variation zero. Hence we cannot ignore the second–order term

$$\frac{1}{2}\int_a^b f''(x)\ (dx)^2$$

in the case that $x = B$, but we can ignore all higher–order terms.

[1] This nonsense can be made precise, promise.

We thus have the following rules for stochastic calculus:

$$(dB_t)^2 = dt$$
$$dB_t \cdot dt = (dt)^2 = 0$$

Suppose that $f(t,x)$ is a $C^{1,2}$–function, and let $X_t = f(t, B_t)$. Applying these rules to a second order Taylor series, we obtain:

Theorem: (Itô's Formula)

$$dX_t = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial B^2}\right) dt + \frac{\partial f}{\partial B}\, dB_t$$

Ordinary calculus shows that for a function $f(t,x)$ we have

$$df = \frac{\partial f}{\partial t}\, dt + \frac{\partial f}{\partial x}\, dx$$

In stochastic calculus, we get another term, due to the non–zero quadratic variation of Brownian motion.

Example 7.2.1 Take our model for stock prices in terms of the returns process:

$$S_t = e^{R_t} \qquad \mathrm{R}_t = \mu t + \sigma B_t$$

Using Itô's formula, we get

$$dS_t = 0\ dt + e^{R_t}\ dR_t + \tfrac{1}{2}e^{R_t}\ (dR_t)^2 = S_t[dR_t + \frac{1}{2}(dR_t)^2]$$

Now

$$dR_t = \mu\ dt + \sigma\ dB_t \qquad (dR_t)^2 = \sigma^2\ dt$$

so

$$dS_t = S_t[\mu\ dt + \sigma\ dB_t + \tfrac{1}{2}\sigma^2\ dt = (\mu + \tfrac{1}{2}\sigma^2)\ dt + \sigma\ dB_t$$

as claimed earlier.

□

Now let's have another look at volatility. The GBM model for stock prices is

$$dS_t = \alpha S_t\, dt + \sigma S_t\, dB_t$$

Thus

$$\mathbb{E}\left[\frac{dS}{S}\right]^2 = \sigma^2\ dt$$

and thus $\sigma^2\ dt$ is the variance of the return of the stock over a small period dt.
It follows that σ is the standard deviation of the annual return of the stock S. This can be measured from market data.

Can we also measure the drift α? This is much harder[2], because over short periods, the dB_t–term dominates the dt–term. The "correct" real-world dynamics of a share price is *difficult to estimate*: We can get the volatility, but not the drift. Amazingly, *we don't care*, as you will see shortly. To calculate option values we need only the volatility, not the drift.

[2]Martin Davis once claimed, in a talk that I attended, that one would need 1500 years of data to get a reasonably accurate estimate of the drift — I'm no statistician, though.

7.3 The Black–Scholes Model

7.3.1 The Black-Scholes PDE

Using Itô's formula, it is not hard to derive a partial differential equation for European style derivatives.

Consider again market with a share S_t whose price process satisfies the SDE

$$dS = \mu S\, dt + \sigma S\, dB_t$$

Let the risk–free interest rate be r, and let A_t be the riskless bank account, with dynamics

$$dA_t = rA_t\, dt$$

Let $V(t, S_t)$ be European–style derivative whose value depends on both the share price and time. Consider a portfolio Π which contains 1 derivative, and n shares, i.e. its value is

$$\Pi_t = V_t + nS_t$$

A small amount of time dt later, the share price has changed. The value of the portfolio changes by

$$d\Pi_t = dV_t + n\, dS_t$$

By Itô's Formula,

$$\begin{aligned} dV_t &= \frac{\partial V}{\partial t}\, dt + \frac{\partial V}{\partial S}\, dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\, dS^2 \\ &= \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S\frac{\partial V}{\partial S}\, dB_t \end{aligned}$$

Hence

$$\begin{aligned} d\Pi_t = &\left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + n\mu S\right) dt \\ &+ \sigma S\left(\frac{\partial V}{\partial S} + n\right) dB_t \end{aligned}$$

Thus

$$\begin{aligned} d\Pi_t = &\left(\frac{\partial V}{\partial t} + \mu S\left[\frac{\partial V}{\partial S} + n\right] + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right) dt \\ &+ \sigma S\left[\frac{\partial V}{\partial S} + n\right] dB_t \end{aligned}$$

Now if we take $n = -\frac{\partial V}{\partial S}$ (i.e. the portfolio is short $-\frac{\partial V}{\partial S}$ shares), then the portfolio is unaffected by the random changes in stock prices:

$$d\Pi_t = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right) dt \tag{7.1}$$

Thus, for a brief moment, the portfolio is risk–free. By a *no–arbitrage argument*, it must *earn the same return as the risk–free bank account*[3], i.e.

$$d\Pi_t = r\Pi_t\, dt = r\left(V - \frac{\partial V}{\partial S}S\right) dt \tag{7.2}$$

Equating (7.1) and (7.2), we get

[3]This is the *crux* of the argument!

$$\boxed{\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0}$$

This is the famous **Black–Scholes** PDE. It is a second–order parabolic PDE, i.e. essentially a heat equation. Most of the PDE's encountered in finance are of a similar type.

Note that if a portfolio contains $\frac{\partial V}{\partial S}$ shares, then the change in the portfolio value is the same as the change in the value of the derivative. The quantity $\frac{\partial V}{\partial S}$ is called the *delta* of the derivative. One can thus *synthetically replicate* any European style derivative with underlying share S by holding, at any time, delta–many shares. This procedure is called *delta hedging.*

Consider a European call option C on a share S with strike K and maturity T. The volatility of the underlying share S is σ and the risk–free rate is r. To find the value of the call option, we must solve the following boundary value problem:

$$\begin{cases} \dfrac{\partial C}{\partial t} + \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 C}{\partial S^2} + rS\dfrac{\partial C}{\partial S} - rC = 0 \\ C(T) = (S_T - K)^+ \end{cases}$$

It is now clear why we don't care about the drift μ of the underlying asset S: It does not appear in the Black–Scholes PDE, and is therefore irrelevant to pricing derivatives.

7.3.2 Pricing in the Risk–Neutral World

In this section we calculate the Black–Scholes prices of vanilla European options. However, we use a slightly subtle probabilistic argument, rather than a brute force "solve the PDE" approach.

In the previous section, we deduced the Black–Scholes PDE for a European–style derivative V:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

Note once again that the *drift* μ *does not occur* in the Black–Scholes PDE, though the volatility σ *does* appear. Hence the price of V is independent of μ, i.e. different values of μ will give the same price.

Thus, for example, the price of a call option on a share S with a given strike K and maturity T will be the same whether μ is very small or very big. This may seem *counterintuitive*, because if μ is very big, it seems as though the option is more likely to expire in–the-money. One would therefore think that the call option price should be an increasing function of μ. But your intuition is just plain wrong.

Since we don't care about the drift rate μ of an underlying asset, we may as well simplify our asset price dynamics by assuming that *all assets have the same drift.* Now the riskless asset (bank account) has drift r, and r occurs in the Black–Scholes PDE. We can't change the drift of the riskfree bank account without changing the PDE, and thus the solution to the pricing problem. So if we want to assume that all assets have the same drift, we have to assume that the drift of all assets is the risk–free rate r.

Mathematically, this corresponds to a change of measure — from a real world, unknowable probability measure $\mathbb{P}$ to a knowable, *risk–neutral measure* $\mathbb{Q}$. In the risk–neutral world, the dynamics of S are

$$dS_t = rS_t\, dt + \sigma S\, dB_t$$

Thus we change the drift of the asset from μ to r.

Why is the world where all assets have the same return called the *risk–neutral* world? Ordinarily, investors are influenced by risk: They weigh up the expected return against the riskiness of an investment. Generally, investors are *risk averse*, which means that they require a premium in order to take on risk. Thus assets with greater riskiness (= volatility) have a higher ("real world") expected payoff than assets with less risk. In the risk–neutral world, investors are indifferent to risk, i.e. they do not require a risk premium. The only thing they care about is *expected return*. Such investors will buy assets with a higher expected return, and sell assets with a lower expected return, regardless of the risk involved. Prices will thus adjust so that all assets have the *same* expected return (in equilibrium). Thus in a world where all investors are risk–neutral, all assets will have the same expected return, i.e. the same expected return as the risk–free bank account.

To summarize, prices in the real– and risk–neutral world are the same. It is just probabilities that are changed. Now we can calculate option prices in the risk–neutral world, because the asset price dynamics are known, and so is the distribution of future stock prices.

Now suppose that we can find a portfolio Π of traded assets which exactly hedges the payoff of a European style derivative V, so that

$$\Pi_T = V_T$$

at the derivative's maturity T. Such a portfolio is called a *replicating portfolio*. By the Law of One Price, therefore, we must have $\Pi_0 = V_0$, where Π_0, V_0 are, respectively the values of the replicating portfolio and the derivative at $t = 0$. Thus:

> *If a derivative has a replicating portfolio, then the value of the derivative equals the value of the replicating portfolio.*

Now in the Black–Scholes model, any European style derivative has a replicating portfolio: A portfolio consisting, at any time, of $\Delta = \frac{\partial V}{\partial S}$ shares will exactly replicate the derivative V (delta hedging).

Π_T and V_T are random variables. But since they are identical, they must have the same expectation, in any world. Since the expected return of all traded assets is r in the risk–neutral world, and since Π consists entirely of traded assets, the expected return of Π is also r:

$$\mathbb{E}_{RN}[\Pi_T] = \Pi_0 e^{rT}$$

where Π_0 is the value of the portfolio at $t = 0$. Now since $\Pi_0 = V_0$ (by the Law of One Price) and $\Pi_T = V_T$ (because Π is a replicating portfolio of V), we see that

$$V_0 = e^{-rT}\mathbb{E}_{RN}[V_T]$$

We have therefore discovered the following procedure for valuing a derivative V.

(i) Assume that the drift of the underlying asset S is r (instead of μ). This moves us from the real world to the risk–neutral world.

(ii) Calculate the expected pay–off of V at maturity in the *risk–neutral world*: $\mathbb{E}_{RN}[V_T]$

(iii) Discount to the present to get the price today:

$$V_0 = e^{-rT}\mathbb{E}_{RN}[V_T]$$

The point is that we can't calculate $\mathbb{E}_{\text{real}}[V_T]$, because we do not know the distribution of the underlying S_T in the real world. However, we *can* calculate $\mathbb{E}_{RN}[V_T]$: Since we know the drift of S_T in the risk–neutral world, we can calculate the distribution of S_T here. This brings us to our next topic.

7.3.3 The Distribution of Asset Prices

We have postulated an asset price model

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dB_t$$

where $\mu = r$ in the risk–neutral world. Consider now the function $Y_t = f(S_t) = \ln S_t$. By Itô's formula,

$$\begin{aligned} dY_t &= \frac{1}{S_t} dS_t - \frac{1}{2}\frac{1}{S_t^2}(dS_t)^2 \\ &= (\mu - \frac{1}{2}\sigma^2)\, dt + \sigma\, dB_t \end{aligned}$$

using $(dB_t)^2 = dt, (dt)^2 = 0 = (dB_t)(dt)$.

So $Y_t = \ln S_t$ follows a *Brownian motion with drift*. We can easily solve this SDE to get

$$Y_T - Y_0 = (\mu - \frac{1}{2}\sigma^2)T + \sigma B_T$$

which implies that Y_T is normally distributed with mean $Y_0 + (\mu - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$:

$$Y_T \sim N\left(Y_0 + (\mu - \frac{1}{2}\sigma^2)T, \sigma^2 T\right)$$

Thus the log of the stock price is normally distributed. We say that stock prices are *lognormally distributed* (in the Black–Scholes model).

Definition 7.3.1 A random variable X is said to be lognormally distributed if and only if the random variable $Y = \ln X$ is normally distributed. Equivalently, if X is normally distributed, then e^X is lognormally distributed.

□

The density function of a lognormal variable

Suppose that $X \sim N(\mu_X, \sigma_X^2)$, so that X has density

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-(x-\mu_X)^2/2\sigma_X^2}$$

Let $Y = e^X$, so that Y is lognormally distributed. Let F_Y, f_Y be, respectively, the distribution and density functions of Y. Then

$$F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}(X \le \ln y) = F_X(\ln y)$$

Differentiating,

$$f_Y(y) = F_Y'(y) = F_X'(\ln y)\frac{1}{y} = \frac{1}{y} f_X(\ln y)$$

Thus:

The density of a lognormal random variable Y is given by

$$f_Y(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi\sigma_X^2}} \dfrac{1}{y} \exp\left[-\dfrac{(\ln y - \mu_X)^2}{2\sigma_X^2}\right] & \text{if } y > 0 \\ 0 & \text{if } y \leq 0 \end{cases}$$

where $Y = \ln X$ and $X \sim N(\mu_X, \sigma_X^2)$. Moreover, the mean μ_Y and variance σ_Y^2 of Y are given by

$$\mu_Y = e^{\mu_X + \frac{1}{2}\sigma_X^2} \qquad \sigma_Y^2 = e^{2\mu_X + \sigma_X^2}[e^{\sigma_X^2} - 1]$$

The statements about the mean and variance of a lognormal random variables are left as straightforward exercises in integration. For example, you must show that

$$\mathbb{E}[Y] = \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_0^\infty \frac{1}{y} e^{-\frac{(\ln y - \mu_x)^2}{2\sigma_x^2}} \, dy = e^{\mu_x + \frac{1}{2}\sigma_x^2}$$

We can now prove the following important fact:

Theorem 7.3.2 *Suppose that Y is lognormally distributed, where $\ln Y \sim N(m, s^2)$. Let K be a positive constant. Then*

$$\begin{aligned} \mathbb{P}(Y \geq K) &= N(d_-) \\ \mathbb{E}[(Y - K)^+] &= \mathbb{E}[Y]N(d_+) - KN(d_-) \end{aligned}$$

where

$$d_\pm = \frac{\ln[\mathbb{E}(Y)/K] \pm \frac{1}{2}s^2}{s}$$

Proof: Since $\ln Y \sim N(m, s^2)$, it follows that if we define $X = \frac{\ln Y - m}{s}$, then $X \sim N(0, 1)$, i.e. X is a standard normal random variable. Clearly

$$\begin{aligned} \mathbb{P}(Y \geq K) &= \mathbb{P}(\ln Y \geq \ln K) \\ &= \mathbb{P}\left(X \geq \frac{\ln K - m}{s}\right) \\ &= 1 - N\left(\frac{\ln K - m}{s}\right) \\ &= N\left(\frac{m - \ln K}{s}\right) \end{aligned}$$

where $N(x)$ is the distribution function of a standard normal random variable, and we used the fact that $1 - N(x) = N(-x)$.

But we know that $\mathbb{E}[Y] = e^{m + \frac{1}{2}s^2}$, so that $m = \ln \mathbb{E}[Y] - \frac{1}{2}s^2$. We thus obtain

$$\mathbb{P}(Y \geq K) = N\left(\frac{\ln \mathbb{E}[Y] - \ln K - \frac{1}{2}s^2}{s}\right) = N(d_-)$$

as required.

Now $\mathbb{E}[(Y-K)^+]$ is an integral which can be split up into two parts. In the first region, $Y \geq K$, so that $(Y-K)^+ = Y-K$ (in that region). In the second region, $Y < K$, so that $(Y-K)^+ = 0$. Thus

$$\mathbb{E}[(Y-K)^+] = \int_K^\infty (y-K)f(y)\,dy$$

where $f(y)$ is the density function of Y. It is simpler to work with X, however, so we change variables: Put $x = \frac{\ln y - m}{s}$. Then $y = e^{sx+m}$, and

$$\mathbb{E}[(Y-K)^+] = \mathbb{E}[(e^{sX+m} - K)^+] = \int_{(\ln K - m)/s}^\infty (e^{sx+m} - K)g(x)\,dx$$

where $g(x)$ is the density of the standard normal random variable X. We can split this up into two integrals:

(1) $\int_{(\ln K-m)/s}^\infty e^{sx+m} g(x)\,dx$

(2) $-K \int_{(\ln K-m)/s}^\infty g(x)\,dx$

We simplify the integrand of the first integral by completing the square:

$$\begin{aligned} e^{sx+m} g(x) &= e^{sx+m} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \\ &= \frac{1}{\sqrt{2\pi}} e^{-(x^2-2sx+s^2)/2} e^{m+s^2/2} \\ &= e^{m+s^2/2} g(x-s) \\ &= \mathbb{E}[Y] g(x-s) \end{aligned}$$

where we used the fact that $\mathbb{E}[Y] = e^{m+s^2/2}$. Thus the first integral becomes

$$\int_{(\ln K-m)/s}^\infty e^{sx+m} g(x)\,dx = \mathbb{E}[Y] \int_{(\ln K-m)/s}^\infty g(x-s)\,dx$$

and the $\int_a^\infty g(x-s)\,dx$ is just the probability that a standard normal random variable is greater than $a-s$, which is $1 - N(a-s) = N(s-a)$. Thus

$$\int_{(\ln K-m)/s}^\infty e^{sx+m} g(x)\,dx = \mathbb{E}[Y] N\left(s - \frac{\ln K - m}{s}\right) = \mathbb{E}[Y] N(d_+)$$

using $m = \ln \mathbb{E}[Y] - s^2/2$.

Similarly, but rather more easily, it can be shown that

$$-K \int_{(\ln K-m)/s}^\infty g(x)\,dx = -KN(d_-)$$

and this completes the proof.

$\square$

The distribution of asset prices

We have

$$\frac{dS_t}{S_t} = \mu \, dt + \sigma \, dB_t$$

which we solved to obtain

$$\ln S_t \sim N\left(\ln S_0 + (\mu - \frac{1}{2}\sigma^2)t, \sigma^2 t\right)$$

Thus $S_t = e^X$, where $X \sim N(\mu_X, \sigma_X^2)$ and

$$\mu_X = \ln S_0 + (\mu - \frac{1}{2}\sigma^2)t$$
$$\sigma_X^2 = \sigma^2 t$$

So the density of S_t is

$$f(S) = \frac{1}{\sqrt{2\pi\sigma^2 t}S} e^{-\frac{[\ln S - (\ln S_0 + (\mu - \frac{1}{2}\sigma^2)t)]^2}{2\sigma^2 t}} \qquad S \geq 0$$

and

$$\mathbb{E}[S_t] = e^{\ln S_0 + \mu t} = S_0 e^{\mu t}$$

Replacing μ with r will give the density of S_t in the risk–neutral world.

Example 7.3.3 Consider a (long) forward contract F on an asset with forward price $K = S_0 e^{rT}$. The payoff of F at T is $S_T - K$. Thus the value of the contract today is

$$F_0 = e^{-rT}\mathbb{E}_{RN}[S_T - K]$$

In the risk–neutral world, the asset price dynamics are given by

$$\frac{dS_t}{S_t} = r \, dt + \sigma \, dB_t \quad \text{i.e.} \quad S_t = S_0 e^{(r - \frac{1}{2}\sigma^2)t + \sigma B_t}$$

Thus $\mathbb{E}_{RN}[S_T] = S_0 e^{rT}$, and so the value of the forward contract is

$$F_0 = e^{-rT}[S_0 e^{rT} - K] = 0$$

which is the *correct* value obtained by the (presumably familiar) static replication argument.

If we had used the "real world" drift μ, however, we would have obtained

$$F_0 = e^{-rT}[S_0 e^{\mu T} - K]$$

and this is incorrect.

□

Exercise 7.3.4 Recall that the value of a forward contract at time t is

$$F_t = [S_t - S_0 e^{rt}]$$

Show that F_t satisfies the BS PDE with boundary condition $F_0 = 0$.

□

By the results in the previous section, we have

$$\begin{aligned}\mathbb{E}S_T &= e^{\mu_X+\frac{1}{2}\sigma_X^2}\\ &= S_0e^{\mu T}\end{aligned}$$

So μ is the expected rate of return of the asset S. In particular, in the risk–neutral world the drift of every traded asset is $\mu = r$, so in the risk–neutral world all assets have the same expected rate of return r (which we already knew).

Also, $\mathbb{E}[(\frac{dS_t}{S_t})^2] = \sigma^2\ dt$, which shows that $\sigma^2\ dt$ is the variance of returns over a period dt. We can thus interpret σ to be the standard deviation of the returns on S over a period of one year.

7.4 Option Pricing: The Black–Scholes Formula

Now that we have the density function of the asset price S_T in the risk–neutral world, we can price practically any European claim V with payoff $\Phi(S_T)$:

$$\begin{aligned}V_0 &= e^{-rT}\mathbb{E}_{RN}[\Phi(S_T)]\\ &= \frac{e^{-rT}}{\sqrt{2\pi\sigma^2T}}\int_0^\infty \Phi(S)e^{-\frac{[\ln S-(\ln S_0+(r-\frac{1}{2}\sigma^2)T)]^2}{2\sigma^2T}}\ \frac{dS}{S}\end{aligned}$$

It is easy to evaluate this integral numerically, using Simpson's method, for example.

Consider next a call option with strike K and maturity T. In this case, $\Phi(S_T) = (S_T - K)^+$. Thus:

$$C_0 = e^{-rT}\mathbb{E}_{RN}[(S_T - K)^+]$$

Now in the risk–neutral world, S_T is lognormally distributed, with $\ln S_T \sim N(\ln S_0 + (r - \frac{1}{2}\sigma^2)T, \sigma^2T)$. By Theorem 7.3.2, therefore,

$$\mathbb{E}_{RN}[(S_T - K)^+] = \mathbb{E}_{RN}[S_T]N(d_+) - KN(d_-)$$

where

$$d_\pm = \frac{\ln[\mathbb{E}(S_T)/K] \pm \frac{1}{2}\sigma^2T}{\sigma\sqrt{T}}$$

But $\mathbb{E}_{RN}[S_T] = S_0e^{rT}$, and thus (remembering to discount):

$$C_0 = S_0N(d_+) - Ke^{-rT}N(d_-)$$

where

$$d_\pm = \frac{\ln\frac{S_0e^{rT}}{K} \pm \frac{1}{2}\sigma^2T}{\sigma\sqrt{T}}$$

and $N(x)$ is the distribution function of a standard normal random variable, i.e.

$$N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-t^2/2}\ dt$$

The normal distribution function $N(x)$ can be determined from tables (see at the back of these notes), or by using a statistical package, e.g. the Excel function NORMSDIST.

Exercise 7.4.1 Verify that the formula for the call option above *is* a solution to the BS PDE. Begin by showing that

$$\begin{aligned}\Delta &: \frac{\partial C}{\partial S} = N(d_+)\\ \Gamma &: \frac{\partial^2 C}{\partial S^2} = \frac{N'(d_+)}{\sigma S\sqrt{T}}\\ \Theta &: \frac{\partial C}{\partial t} = -\frac{\sigma S N'(d_+)}{2\sqrt{T}} - rKe^{-rT}N(d_+)\end{aligned}$$

□

The partial derivatives in the above exercise are known as *the Greeks*, and are measures of the sensitivity of an option to its parameters. Other Greeks are

$$\begin{aligned}\rho &: \frac{\partial C}{\partial r}\\ \text{Vega} &: \frac{\partial C}{\partial \sigma}\end{aligned}$$

We can obtain the formula of a put option in the same way that we derived the formula for a call option, i.e. via a long and complicated chain of integrations. However, it is more intelligent to use *put–call parity:*

$$\begin{aligned}P_0 &= C_0 + Ke^{-rT} - S_0\\ &= S_0[N(d_+) - 1] + Ke^{-rT}[1 - N(d_-)]\\ &= -S_0 N(-d_+) + Ke^{-rT}N(-d_-)\end{aligned}$$

i.e.

$$\boxed{P_0 = -S_0 N(-d_+) + Ke^{-rT}N(-d_-)}$$

Finally, as a curiosity, we mention *binary* or *digital options*:

Definition 7.4.2

- A binary call on S with strike K will pay one unit of currency if $S_T \geq K$ at expiry, and nothing otherwise.
- A binary put will pay 1 if $S_T < K$, and nothing otherwise.

□

Let B_c be a binary call with strike K. The boundary value problem for B_c is

$$\begin{cases} \dfrac{\partial B_c}{\partial t} + \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 B_c}{\partial S^2} + rS\dfrac{\partial B_c}{\partial S} - rB_c = 0\\ B_c(T) = I_{\{S_T \geq K\}}\end{cases}$$

Exercise 7.4.3 (1) The solution is of this BVP is given by

$$B_c(0) = e^{-rT} N(d_-)$$

where as before

$$d_- = \frac{\ln \frac{S_0}{K} + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

(2) You should be able to obtain a put–call parity for binary options, and then deduce the value of a binary put from that.

□

7.5 Options on Dividend–Paying Stocks, Futures and Currencies

In this section we discuss the valuation of options on a dividend–paying asset. We consider two alternatives: (i) The case where we know how much will be paid, and when, and (ii) the case where the dividend yield is known, and assumed constant. We shall then see that options on currencies and futures can be regarded as a special case of (ii).

7.5.1 Options on a Dividend–Paying Asset

Consider once more the Black–Scholes formula for a call option:

$$C_0 = e^{-rT} \left[\mathbb{E}[S_T] N(d_+) - K N(d_-) \right]$$

where

$$d_+ = \frac{\ln[\mathbb{E}(S_T)/K] + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}$$

$$d_- = \frac{\ln[\mathbb{E}(S_T)/K] - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} = d_+ - \sigma\sqrt{T}$$

and expectations are taken under the risk–neutral measure.

We will consider two possibilities, namely that a known dividend is paid, or that the share has a known dividend yield. In both cases, we must find $\mathbb{E}[S_T]$, and determine that S_T is lognormally distributed, so that we can apply Theorem 7.3.2 to find call option prices.

Options on a stock paying a known dividend

Suppose that we know that a dividend will be paid out during the life of the option, and let $\mathrm{PV}(D)$ be the $t = 0$–value of this dividend. Then clearly $\mathbb{E}[S_T] = (S_0 - \mathrm{PV}(D))e^{rT}$. When we plug this into the Black–Scholes formula for a call option, we see that

$$C_0 = (S_0 - \mathrm{PV}(D))N(d_+) - Ke^{-rT} N(d_-)$$

where

$$d_+ = \frac{\ln[(S_0 - \mathrm{PV}(D))/K] + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

$$d_- = d_+ - \sigma\sqrt{T}$$

This formula can also be justified as follows: Once it's known that a certain dividend is going to be paid at a certain time t_d before maturity T, it is possible to decompose the asset price process into two terms:

$$S_t = S_t^{\text{risky}} + S_t^{\text{riskless}}$$

The riskless component corresponds to the dividend process, i.e.

$$S_t^{\text{riskless}} = \begin{cases} \text{PV}(D)e^{rt} \text{ if } 0 \leq t \leq< t_d \\ 0 \text{ if } t_d \leq t \leq T \end{cases}$$

Here we assume that the date that the dividend is paid out is the same as the date when the share goes ex–dividend. Some small but obvious adjustments must be made when, as is usual, this is not the case.

Note that $S_0^{\text{risky}} = S_0 - \text{PV}(D)$.

We may assume that the risky part S_t^{risky} follows a geometric Brownian motion. This is a modelling assumption. It then follows that $S_T = S_T^{\text{risky}}$ is lognormally distributed, so that we may apply Theorem 7.3.2 to obtain the price of the option. For that, we have to know $\mathbb{E}[(S_T)]$. In the risk–neutral world, however, both S_t^{risky} and S_t^{riskless} have drift r, so it is clear that

$$\begin{aligned} \mathbb{E}[S_T] &= \mathbb{E}[S_T^{\text{risky}}] \\ &= S_0^{\text{risky}} e^{rT} \\ &= (S_0 - \text{PV}(D))e^{rT} \end{aligned}$$

One must be careful about volatility, however. To estimate volatility from market data, one must keep in mind that the risk lies all in S_t^{risky}. Thus volatility must be estimated from the continuous process S_t^{risky}, and not from the discontinuous process S_T. In practice, this is important only if the dividend forms a fairly large proportion of the share price.

The put–call parity formula obtained in Exercise **??** can be used to find the corresponding price for a put option on a dividend–paying stock.

Options on a stock with a known dividend–yield

Suppose that we know that the dividend yield of a stock S is q, so that the stock pays a dividend $qS_t\,dt$ between times t and $t + dt$. Then if

$$dS_t = \mu S_t\,dt + \sigma\,dB_t$$

are the dynamics of the stock without a dividend,

$$dS_t = (\mu - q)S_t\,dt + \sigma\,dB_t$$

are the stock price dynamics *with* dividend. In the risk–neutral world, $\mu = r$, the riskless rate.

We may now apply the Black–Scholes argument to a portfolio Π consisting of one call and n shares: $\Pi = C + nS$. Taking differentials and applying Itô's formula, we get

$$\begin{aligned} d\Pi &= dC + n\,dS + nqS\,dt \\ &= \frac{\partial C}{\partial t}\,dt + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\,dS^2 + \left(n + \frac{\partial C}{\partial S}\right)\,dS + nqS\,dt \end{aligned}$$

As before, setting $n = -\frac{\partial C}{\partial S}$ makes the portfolio Π locally riskless. By an arbitrage argument, it must therefore earn a return equal to the riskless rate r, i.e.

$$d\Pi = r\Pi \, dt \quad \text{when} \quad n = -\frac{\partial C}{\partial S}$$

Comparing the two expressions for $d\Pi$ we obtain a Black–Scholes PDE for an asset paying a continuous dividend:

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + (r-q)S\frac{\partial C}{\partial S} - rC = 0$$

We can now solve as before: Since μ does not occur in the PDE, we may as well assume that all traded assets have the same rate of return, namely the riskless rate. This moves us to the risk–neutral world. Here the drift of an asset S_t with dividend yield q is $r-q$, and S_T will be lognormally distributed, because its underlying process is modelled by a geometric Brownian motion. The only thing we need to know in order to apply Theorem 7.3.2 is $\mathbb{E}_{RN}[S_T]$. Now since $\mathbb{E}[S_T] = S_0 e^{rT}$ when the drift of S_t is r, it follows that

$$\mathbb{E}_{RN}[S_T] = S_0 e^{(r-q)T}$$

when the drift is $r-q$, i.e. when the dividend yield is q.

It now follows easily that

$$C_0 = S_0 e^{-qT} N(d_+) - K e^{-rT} N(d_-)$$

where

$$d_+ = \frac{\ln[S_0/K] + (r - q + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$
$$d_- = d_+ - \sigma\sqrt{T}$$

The put–call parity relationship for options on dividend–paying assets obtained in Exercise **??** can now be used to obtain the price of a put on an asset with a continuous dividend yield.

To summarize: To value an option on an asset with

- Known dividends: Replace S_0 by $S_0 - \mathrm{PV}(D)$ in the standard Black–Scholes formula;
- Known dividend yield: Replace S_0 by $S_0 e^{-qT}$ in the standard Black–Scholes formula.

7.5.2 Options on Futures

We derive in this section Black's model for futures option prices. A futures option is an option to acquire or deliver a futures contract, for example a call option with strike K and maturity T on a futures contract with maturity $T^* \geq T$. If $F(t, T^*)$ is the futures price at time t, then the payoff of the option is

$$\text{Payoff } = (F(T, T^*) - K)^+$$

Suppose that the futures contract is on an underlying asset that follows a geometric Brownian motion: $dS_t = \mu S_t \, dt + \sigma S_t \, dB_t$. We assume that interest rates are constant, so that the theoretical futures price is equal to the forward price: $F(t, T^*) = S_t e^{r(T^*-t)}$. Applying Itô's formula, we can easily calculate

$$\begin{aligned} dF_t &= e^{r(T^*-t)} \left[dS_t - rS_t \, dt \right] \\ &= (\mu - r) \, dt + \sigma \, dB_t \end{aligned}$$

In the risk–neutral world, $\mu = r$, so that

$$dF_t = \sigma F_t \, dB_t$$

Thus the futures price has no drift in the risk–neutral world: It is a *martingale*.

Note that the volatility of the futures price and the underlying stock are the same.

Now the stochastic differential equation $dF_t = \sigma F_t \, dB_t$ is the same as that satisfied by an asset that pays a continuous dividend yield of $q = r$ — We saw that the drift of an asset with dividend yield q is $r - q$ in the risk–neutral world), so that an asset with drift 0 must have a dividend yield of r. It is thus *as if* a futures price is an asset with dividend yield r.

We can therefore apply the results of the previous section to calculate the price of a futures call option. It is

$$C_0 = e^{-rT}[F(0,T^*)N(d_+) - KN(d_-)]$$

where

$$d_+ = \frac{\ln[F(0,T^*)/K] + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}$$
$$d_+ = d_- - \sigma\sqrt{T}$$

As usual, you can figure out the put price yourself.

In South Africa, futures options are subject to margin calls, which changes Black's formula slightly.

7.5.3 Options on Currencies

A currency option is the option to buy or sell a preagreed amount of foreign currency at a predetermined exchange rate at a preagreed future time. For example, suppose that a US company has the option to buy £1 million for \$K at time T. This amounts to a call option on the £. Let S_t be the foreign/local (e.g. £/\$) exchange rate at time t, e.g. £1 is an asset with price S_t. The above option has payoff at maturity

$$\text{Payoff } = (S_T - K)^+$$

Let r_f be the foreign riskless rate, and let r be the local riskless rate. Assume that the exchange rate S_t follows a geometric Brownian motion. Now a pound is an asset which can be invested in a UK bank, where it will pay a continuous dividend yield of r_f. We may therefore regard the exchange rate as an asset with a continuous dividend yield r_f. In the risk–neutral world, therefore, the exchange rate has the following dynamics:

$$dS_t = (r - r_f)S_t \, dt + \sigma S_t \, dB_t$$

where the *volatility* σ is a number defined by the property

$$\text{Standard deviation of } \ln S_T = \sigma\sqrt{T}$$

Thus S_T is lognormally distributed, and we see that the price of a foreign currency call option is

$$C_0 = S_0 e^{-r_f T} N(d_+) - Ke^{-rT} N(d_-)$$

where

$$d_+ = \frac{\ln[S_0/K] + (r - r_f + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$
$$d_- = d_+ - \sigma\sqrt{T}$$

Chapter 8

Hedging with Options: The Greeks

The problem of hedging cropped up on a number of occasions in previous chapters. We discussed hedging with futures in Chapter 3, and addressed the problem of hedging interest rate risk in Chapters 5 and 6. Then, in Chapter 7, we derived the Black–Scholes formula using the following idea:

> *The value of any derivative security is equal to the value of any portfolio that perfectly hedges it.*

This idea was used much earlier, e.g. to determine forward prices, forward rates and swap rates. In Chapter 7, however, we used a portfolio that was changing in continuous–time to dynamically hedge an option. Now continuous–time hedging is impossible: You cannot trade at every instant in time. Even if it *were* possible, it would be impractical: Transaction costs would be infinite. The Black–Scholes formula is therefore accurate only to the extent that we can hedge options in real time. The aim of this chapter is to show how to do this.

8.1 Introduction

Let V be a financial instrument of portfolio based on an underlying asset S, e.g. an option. Throughout this chapter, we will use the following notation:

- S = current spot price;
- σ = volatility;
- r = c.c. risk–free rate;
- T = time to maturity;
- Price of option: $V = V(S, r, \sigma, T)$

(For brevity's sake we shall refer to V as an *option* throughout this chapter, but bear in mind that V could be any derivative or portfolio based on V).:

Also define the following quantities, known as the Greeks[1]:

[1] They're called "the Greeks" because their names are letters of the Greek alphabet, all except $\mathcal{V}$, which is pronounced "vega". How vega got to be a Greek is a mystery: Some claim that it was named after a car. Just possibly, however, that a slightly clueless trader once looked at a star chart, saw *Alpha* and *Beta* Centauri just 4 light years away, and saw Vega not far off. Vega is three times the size of the sun, 50 times as bright, and just 26 light years away, but its name is not in the Greek alphabet.

$$\Theta = \frac{\partial V}{\partial t}$$
$$\Delta = \frac{\partial V}{\partial S}$$
$$\Gamma = \frac{\partial^2 V}{\partial S^2}$$
$$\mathcal{V} = \frac{\partial V}{\partial \sigma}$$
$$\rho = \frac{\partial V}{\partial \rho}$$

Note that the Black–Scholes PDE contains the top three Greeks:

$$\Theta + rS\Delta + \frac{1}{2}\sigma^2 S^2 \Gamma - rV = 0$$

8.2 Delta

The *delta* of V is simply $\Delta = \frac{\partial V}{\partial S}$. Graphically, Δ is the *slope* of the curve V vs S. Thus if $\Delta = 0.2$, then an increase of \$0.01 in the price of the underlying asset will result in an increase of approximately \$0.002 in the price of the option. It follows that delta measures the sensitivity of the option to changes in the underlying asset.

Since $\frac{\partial}{\partial S}$ is a linear operator, the delta of a portfolio is the sum of the deltas of its constituents. (This is clearly true for all the Greeks.) We made use of the linearity of Δ in the Black–Scholes argument: In the derivation of the Black–Scholes PDE, we started with a portfolio Π consisting 1 derivative V and n shares S. The portfolio Π, the derivative V and the underlying share all have their own delta. Clearly $\Delta_S = 1$, so that $\Delta_\Pi = \Delta_V + n$ by linearity. Black and Scholes now argue as followes: If we set $n = -\Delta_V$, then $\Delta_\Pi = 0$, so that the portfolio Π is insensitive to changes in the underlying, i.e. it is riskless. Hence it must earn the risk–free rate of return, i.e.

$$d\Pi = r\Pi\, dt$$

Compare this with the expression obtained by Ito's formula,

$$d\Pi = \frac{\partial \Pi}{\partial t}\, dt + \frac{\partial \Pi}{\partial S}\, dS + \frac{1}{2}\frac{\partial^2 \Pi}{\partial S^2}\, dS^2$$

and the Black–Scholes PDE pops out. Thus Δ_V is the number of shares that a portfolio of V and S should be *short* in order to be instantaneously riskless.

However, Δ is not a static quantity, but changes over time. In order to *delta–hedge* a portfolio, the delta must be calculated frequently, and the portfolio must be rebalanced frequently: This is called *dynamic hedging*.

Delta hedging protects a portfolio from losses (and gains) due to movements in the price of the underlying. If a portfolio Π is *delta–neutral* (i.e. has $\Delta = 0$), then it will not be affected by small instantaneous changes in the value of the underlying, because

$$\delta\Pi \approx \frac{\partial \Pi}{\partial S}\delta S = 0$$

Delta for "Linear" Securities

It is clear that the Δ of a share S (with respect to itself) is $\frac{\partial S}{\partial S} = 1$. To put this differently, to hedge a portfolio of one share, you should short one share.

Since the value of a long forward contract on S with delivery price K and maturity T is just

$$F = S - Ke^{-rT}$$

(assuming that it pays no dividends during the life of the forward contract) it is clear that a long forward contract on S also has $\Delta = 1$. A short forward contract has $\Delta = -1$.

Exercise 8.2.1 Calculate the Δ of a long forward on S given that

(a) S is known to pay a dividend of D at a time t_d before maturity T.

(b) S has a continuous dividend yield of q.

□

Although futures contracts are similar to forward contracts, they are not the same. Strikingly, the Δ of a future and a forward contract are very quite different, even if interest rates are deterministic, or constant. Suppose that we are long a futures contract on S. For simplicity, assume that the riskless rate r is constant, so that the futures price is $\mathcal{F} = Se^{rT}$. If the value of the underlying increases by δS, then the futures price changes by $\delta F = e^{rT}\delta S$. Because of the marking–to–market process, this results in an *immediate gain* of $e^{rT}\delta S$ — for a forward contract, this gain would only have been realized at maturity. It follows that the Δ of a long futures contract (on a non–dividend paying asset) is $\Delta = e^{rT}$.

Exercise 8.2.2 Calculate the Δ of a long futures contract on an asset with continuous dividend yield q (assuming that the riskless rate of interest is constant).

□

Example 8.2.3 Because futures are very liquid and have zero initial cost, they are often used to delta hedge a portfolio. Suppose, for example, that a trader manages a portfolio of shares and options on the share. The portfolio is due to be closed out or sold at time T, but the trader has already made a neat profit on this portfolio, and would like to lock these in, i.e. the trader wants to hedge against portfolio losses due to further movements in the underlying. Suppose that the Δ of the combined portfolio is 1000. Thus every small move δS in the underlying causes a move in the value of the portfolio that is 1000 times bigger. In order to hedge, the trader can, at zero cost, add just enough futures to the portfolio to make the portfolio Δ equal to zero. If every future is on 100 shares, for example, then the trader needs to short $\frac{1000}{100e^{rT}}$ futures with maturity T. If futures with maturity T are not available, then the trader can accomplish the same by shorting $\frac{1000}{100e^{rT^*}}$ futures with maturity T^*.

□

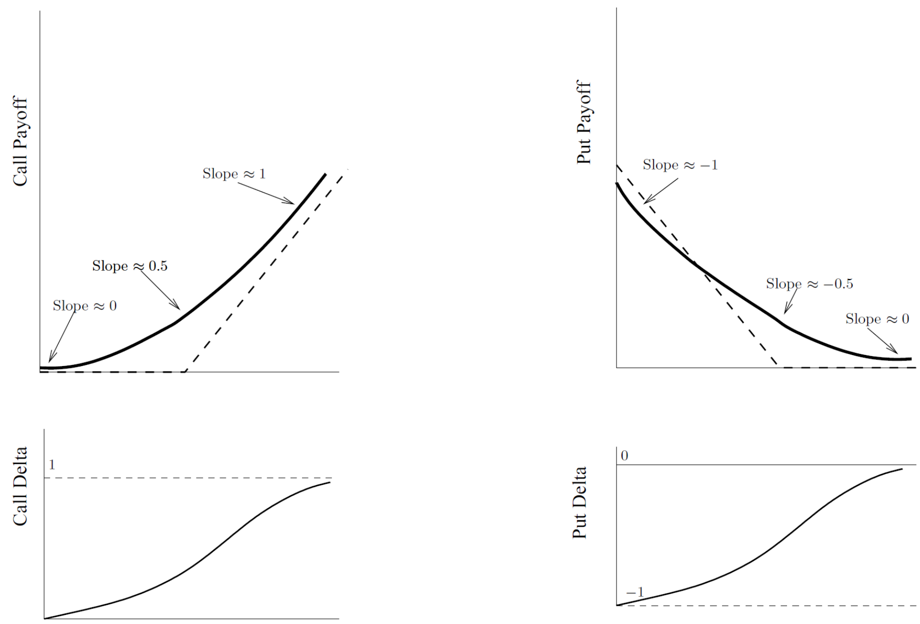

Figure 8.1: Behaviour of *delta* for vanilla options on a non–dividend paying asset.

Delta of Vanilla Options

Let C and P be, respectively, the price of a call and a put option with strike K and maturity T on a share with dividend yield q. It is left as an exercise to show that

$$\Delta_C = e^{-qT} N(d_+) \qquad \Delta_P = e^{-qT}[N(d_+) - 1]$$

where

$$d_+ = \frac{\ln\frac{S}{K} + (r - q + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

(First derive the result for a call, and then use put–call parity to derive the result for a put.)

The above formulas for Δ also hold for options on assets which behave like a dividend–paying asset: For foreign currencies, put $q = r_f$, the foreign riskless rate, and for futures put $q = r$.

Since $0 \leq N(d_+) \leq 1$, the delta of a call is positive, whereas the delta of a put is negative. This is obvious from the graphs of call– and put prices versus S: The graph of a call is upwards sloping, and the graph of a put is downwards sloping. It also makes financial sense: A call becomes more valuable as the underlying asset price increases, whereas a put becomes more valuable if the underlying asset price decreases.

Consider now vanilla options on a non–dividend paying asset. From the graphs, it is clear that delta increases from 0 to 1 for a call, and from-1 to 0 for a put. Thus the deltas of both calls and puts are increasing functions of the underlying asset.

For vanilla call options, $\Delta > 0$, and for put options, $\Delta < 0$. This reflects the fact that the writer of a call hedges by buying shares, whereas the writer of a put must short shares in order to hedge. Note that Δ is an increasing function of S, for both calls and puts.

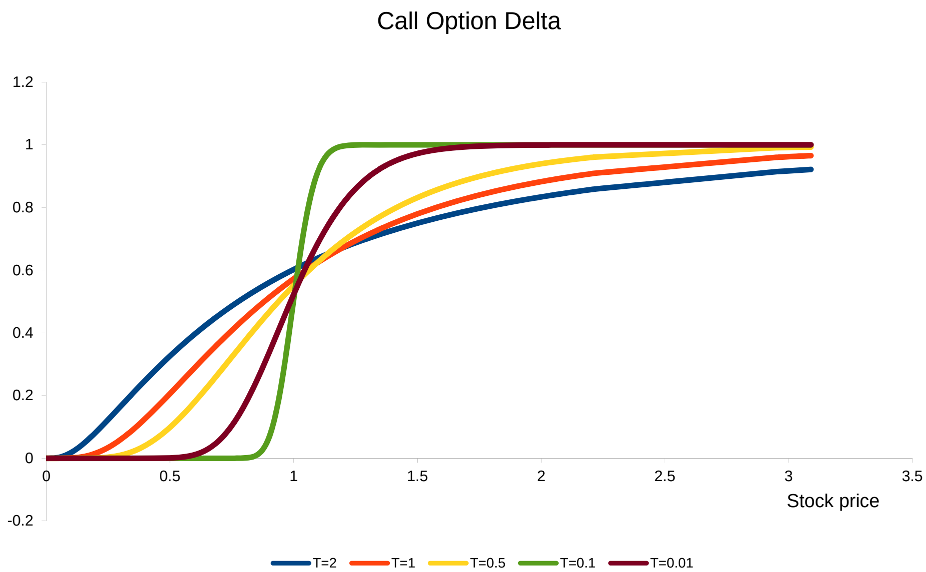

Figure 8.2: Behaviour of delta for call options over time.

Deep–in–the money calls have $\Delta \approx 1$. This reflects the fact that it is almost certain that a deep–in–the—money call will be exercised. Thus an increase of δS in the stock price leads to an increase of δS in the payoff of the call. Similarly deep–out–of–the–money calls have $\Delta \approx 0$, because if the stock price increases by δS, the payoff of the call will still be zero (being still unlikely to be exercised).

For at–the–money calls, $\Delta \approx 0.5$. Deep–in–the–money puts have $\Delta \approx -1$, deep–out–of–the–money puts have $\Delta \approx 0$, and at–the–money puts have $\Delta \approx -0.5$.

Next, we investigate what happens to Δ over time. Suppose that today the share price is S and that we own a three–month call with strike K and a delta of Δ_3, and a one–month call, with the same strike and a delta of Δ_1. How are Δ_3 and Δ_1 related? It depends: The deltas of long–term options are less curved than those of short–term options. We will have $\Delta_3 \geq \Delta_1$ if the calls are out–of–the–money, whereas $\Delta_3 \leq \Delta_1$ if the calls are in–the–money. This behaviour is illustrated in Figure 8.2

It is clear that as the time to maturity approaches zero, the delta becomes steeper around the strike. An instant before maturity, the delta of a call is zero if the call is out of the money, and one if the call is in the money. For at–the–money calls, the delta is 0.5 an instant before maturity. We can see this not only from the graphs, but also via direct calculation. Since $\Delta_C = N(d_-)$, we see that:

$$\lim_{T \downarrow 0} N(d_+) = \lim_{T \downarrow 0} N\left(\frac{\ln(S/K)}{\sigma\sqrt{T}}\right)$$

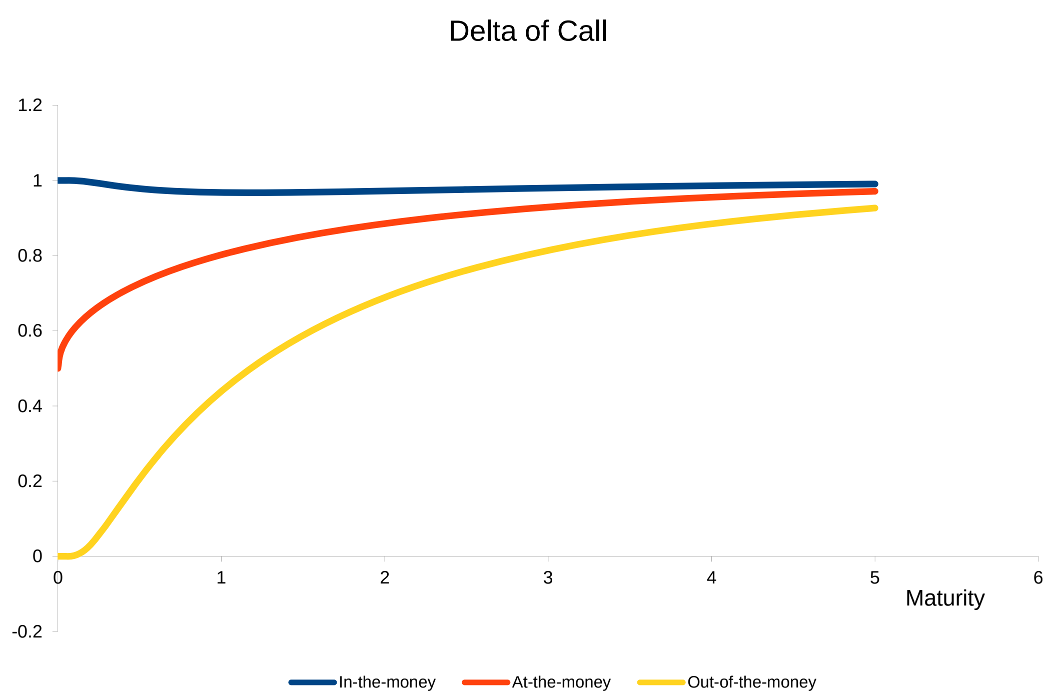

Figure 8.3: Behaviour of delta for call options as a function of maturity.

But $\ln(S/K)$ is: (i) > 0 if $S > K$, (ii) $= 0$ if $S = K$, and (iii) < 0 if $S < K$, so that as $T \downarrow 0$ we obtain

$$\Delta_C = \begin{cases} N(+\infty) = 1 \text{ if } S > K \\ N(0) = 0.5 \text{ if } S = K \\ N(-\infty) = 0 \text{ if } S < K \end{cases}$$

On the other hand,

$$d_+ = \frac{\ln(S/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \to \infty \quad \text{as } T \to \infty$$

Hence

$$\Delta_C \to 1 \qquad \text{as} \qquad T \to \infty$$

We can therefore plot Δ_C as a function of maturity T — see Figure 8.2.

Delta Hedging

In the derivation of the Black–Scholes PDE, we started with a portfolio consisting of one option and a number of shares. We noted that if at any time our portfolio is short Δ shares, then the portfolio is *instantaneously riskless*.

To exclude arbitrage, therefore, our portfolio must earn the same rate of return as the riskless rate.

It is therefore possible to create a portfolio of shares Π which has the same response to changes in the underlying as does the option V:

$$d\Pi = \Delta_\Pi \, dS \qquad dV = \frac{\partial V}{\partial S} \, dS \quad \text{so } d\Pi = dV$$

Such a portfolio will replicate the payoff of the option, and is therefore called a replicating portfolio. The replicating portfolio of an option is, in essence, a *synthetic option*: Financial institutions write options, and then hedge these options by maintaining an equivalent position in the underlying share. This process is called (dynamic) delta hedging; it is dynamic, because Δ is not constant over the life of the option, and the replicating portfolio must be rebalanced frequently if it is to be an effective hedge. The greater the frequency of rebalancing, the more accurate the hedge. (In the derivation of the Black–Scholes PDE, we assumed that the portfolio was rebalanced at every instant.)

Pension funds are typically allowed to make only very restricted use of derivatives (or none at all). If fund managers want to lock in on profits, but are not legally allowed to buy put options, they can create a synthetic put instead.

We show here how to synthetically hedge a call option using delta hedging. Suppose that on 21 March 2003, a trader writes a one–year call with strike $K = 1.00$ on a share S. Today's stock price is $S_0 = 0.95$. The stock volatility is 40% and the riskless rate is 6% (c.c). The Black–Scholes price of this option is therefore \$0.1537, as you can easily verify. The trader is short a call option, and will hedge the option by taking a long position in shares. The trader wants to follow a hedging strategy which will ensure that she owns one share if the call expires in–the–money, and no shares if the call expires out–of–the–money. Moreover, the cost of hedging must be (approximately) equal to the Black–Scholes price of the option — i.e. the cost of hedging must be covered by the premium she receives. The trader decides to rebalance her portfolio every 15 days. For simplicity, we assume a 360–day year, so that the portfolio is readjusted 24 times during the year. Here is what she does: Every 15 days, the trader will observe the stock price in the market, and then calculate the Δ of a call with expiry 21 March 2004. She will then buy or sell enough shares to ensure that she owns exactly Δ shares at each rebalancing date.

For example, on 21 March, the Δ of the option is 0.5878. The trader therefore buys 0.5878 shares, for a total of $0.95 \times 0.5878 = \$0.5584$. She finances this purchase by borrowing the cost at 6% c.c for 1 year. Thus at maturity, she will owe \$0.5929.

15 days later the share price is 1.017. The delta of the option is now 0.6502. The trader needs an additional $0.6502 - 0.5878 = 0.0624$ shares. These are purchased at a cost of $0.0624 \times 1.017 = \$0.0635$. This purchase is financed by borrowing the cost at 6% (c.c.) for 345 days. So at maturity, she will owe an additional \$0.0674.

She repeats this procedure every 15 days. Sometimes she will need to sell shares, not buy them. She then invests the money received from the sale at 6% until maturity.

The following table shows a possible strategy that the trader might follow:

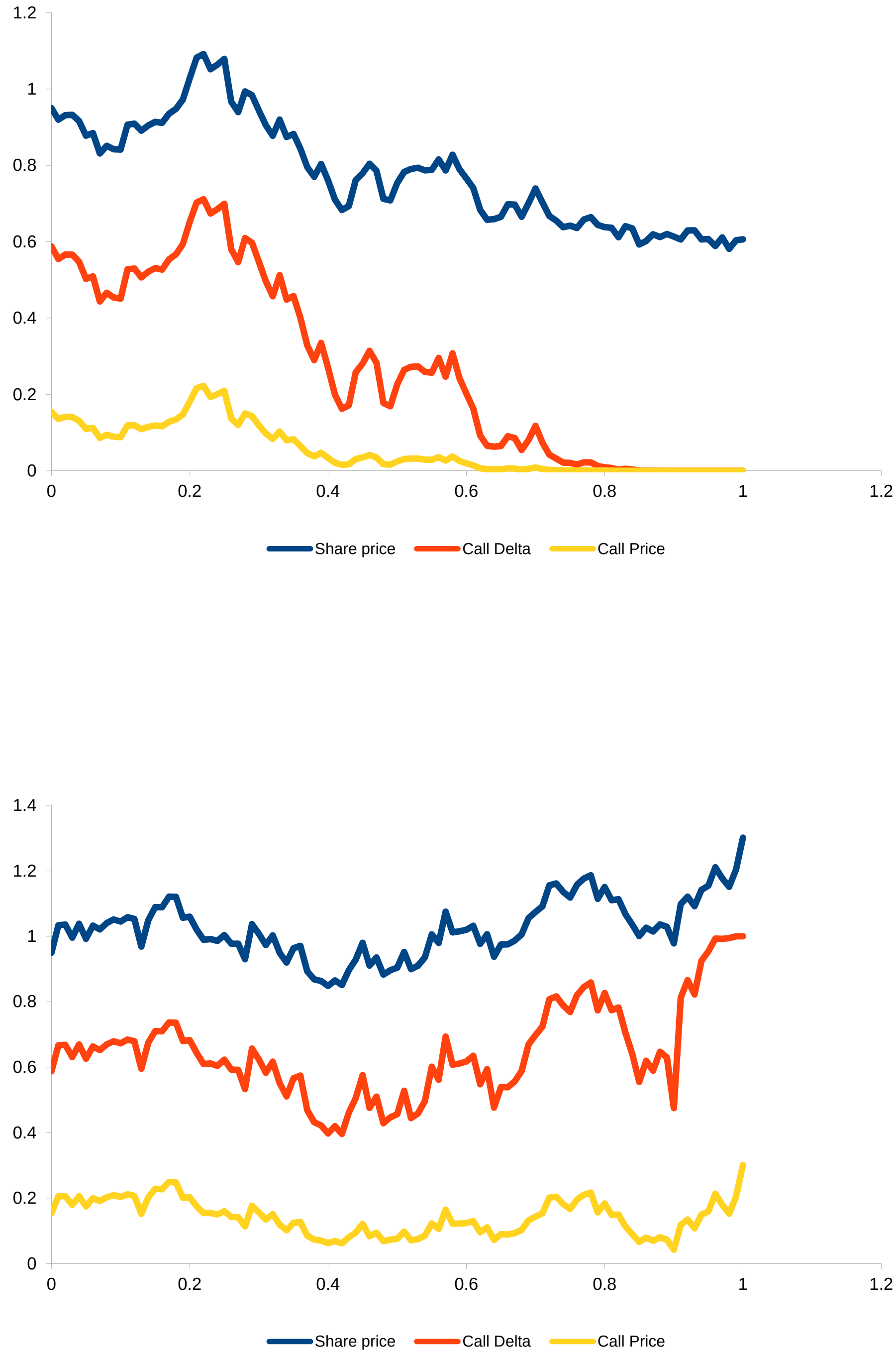

Figure 8.4: Share price, option price and option delta for a $K = 1$, $T = 1$ call option

Date (15 days)	Share Price	Delta	Cost (FV)
0	0.9500	0.5878	0.5929
1	1.0171	0.6502	0.0674
2	1.0889	0.7114	0.0706
3	1.2159	0.8023	0.1168
4	1.3018	0.8512	0.0671
5	1.2282	0.8130	-0.0490
6	1.1109	0.7280	-0.0990
7	1.0700	0.6900	-0.0425
8	0.9888	0.5992	-0.0937
9	0.9329	0.5227	-0.0743
10	0.8801	0.4401	-0.0755
11	0.9033	0.4651	0.0234
12	1.0975	0.7178	0.2865
13	1.1265	0.7507	0.0382
14	1.1317	0.7594	0.0101
15	1.1378	0.7707	0.0132
16	1.1196	0.7553	-0.0176
17	1.0339	0.6345	-0.1275
18	0.9966	0.5627	-0.0728
19	0.8642	0.2612	-0.2645
20	0.7817	0.0861	-0.1386
21	0.7692	0.0417	-0.0345
22	0.7257	0.0037	-0.0277
23	0.6847	0.0000	-0.0026
24	0.6738	0.0000	0.0000

The column on the right shows the cost of each rebalancing act, adjusted for interest. Adding up the entries in this column, we see that the total future value (at maturity) of the hedging costs are \$0.1661. The present value of this is $\$0.1661e^{-0.06} = 0.1564$. This compares quite well with the Black–Scholes price, \$0.1541, of the option, even though the replicating portfolio was rebalanced only 24 times. If the portfolio is rebalanced on a daily basis, the cost of hedging and the cost of of the option would be virtually the same. It is a very instructive exercise to create a spreadsheet which calculates the cost delta hedging for different rebalancing frequencies, and compare these to the Black–Scholes price.

Note that in the above example, the call expired out–of–the–money. At maturity, the table shows that Δ of the call is 0.0000. Thus at maturity, the trader owns zero shares. And she doesn't need to own any, because the call expired out–of–the–money.

Under a different scenario, the call might have expired in–the–money. The delta hedging procedure would then have guaranteed that the trader owns exactly one share at maturity. The future value of the cost of the hedge (i.e. the sum of the entries in the rightmost column) would then have been approximately \$1.164, but the trader would receive the strike $K = 1$ at maturity, so that the total future value of the cost would be about \$0.164. The present value of this is, again, about \$0.154.

In reality, every rebalancing act is subject to transaction costs. Even if it were feasible to exactly hedge a derivative by delta–hedging at every instant (as in the derivation of the Black–

Scholes PDE), the transaction costs would then be infinite. Thus a hedger must compromise betwen accuracy and cost. For example, while it would be too expensive to delta–hedge each derivative at the end of every trading day, it may well be feasible to hedge an entire portfolio of such derivatives — only a single transaction will take care of the whole lot. And by watching gamma as well as delta, a trader can ensure that a portfolio does not have to be rebalanced too frequently.

8.3 Gamma

The Γ of an option or a portfolio is the second derivative of its value with respect to the price of the underlying:

$$\Gamma = \frac{\partial^2 V}{\partial S^2}$$

Thus Γ captures the second–order effects in price changes:

$$dV \approx \frac{\partial V}{\partial S}\, dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\, dS^2 = \Delta_V \, dS + \frac{1}{2}\Gamma_V \, dS^2$$

This should remind you of the corresponding notions for bond prices as a function of yield:

$$\frac{dP}{P} \approx -D^* \, dy + \frac{1}{2} C \, dy^2$$

Thus Δ is like duration, whereas Γ is like convexity.

Like convexity measure the curvature of the bond price as a function of yield, Γ measures the curvature of the value of an option V as a function of the price of the underlying S:

- If Γ is positive, then the graph of $V = f(S)$ is concave up, whereas if Γ is negative, it is concave down.
- If $|\Gamma|$ is large, then the graph of $V = f(S)$ is highly curved.

Thus if Γ is large, the second order effects significantly affect the price change.

Γ plays an important role in delta hedging an option: Since $\Gamma = \frac{\partial \Delta}{\partial S}$ is just the rate of change of Δ with respect to the underlying, a large Γ implies that Δ is changing very rapidly. It follows that the portfolio must be rebalanced more frequently (to remain delta–hedged) if $|\Gamma|$ is large. This is because delta hedging protects against only first order moves in the underlying asset. As $|\Gamma|$ increases, second order effects become important as well.

Just as a bond with greater convexity is preferable to a bond with smaller convexity (given that the yield and duration are the same), a derivative with larger Γ is preferable to one with smaller Γ, and for very similar reasons. Figure 8.3 shows why. If $|\Gamma|$ is large, then the price curve of a derivative (as a function of the underlying S) moves away from the tangent line more rapidly than if $|\Gamma|$ is small. This is beneficial if Γ is positive, but adversely affects the option price if Γ is negative. To see this, first suppose that Δ and Γ are positive (e.g. for a long call). If S increases, then the value of the derivative increases more rapidly than it would if the price function is uncurved. Similarly if S decreases, then the derivative loses less value than it would if the price function is uncurved. Thus whether S increases or decreases, the derivative's price function behaves better if it is curved than if it is uncurved.
Next suppose that Δ is negative and Γ is positive (e.g. for a short put). If S increases,

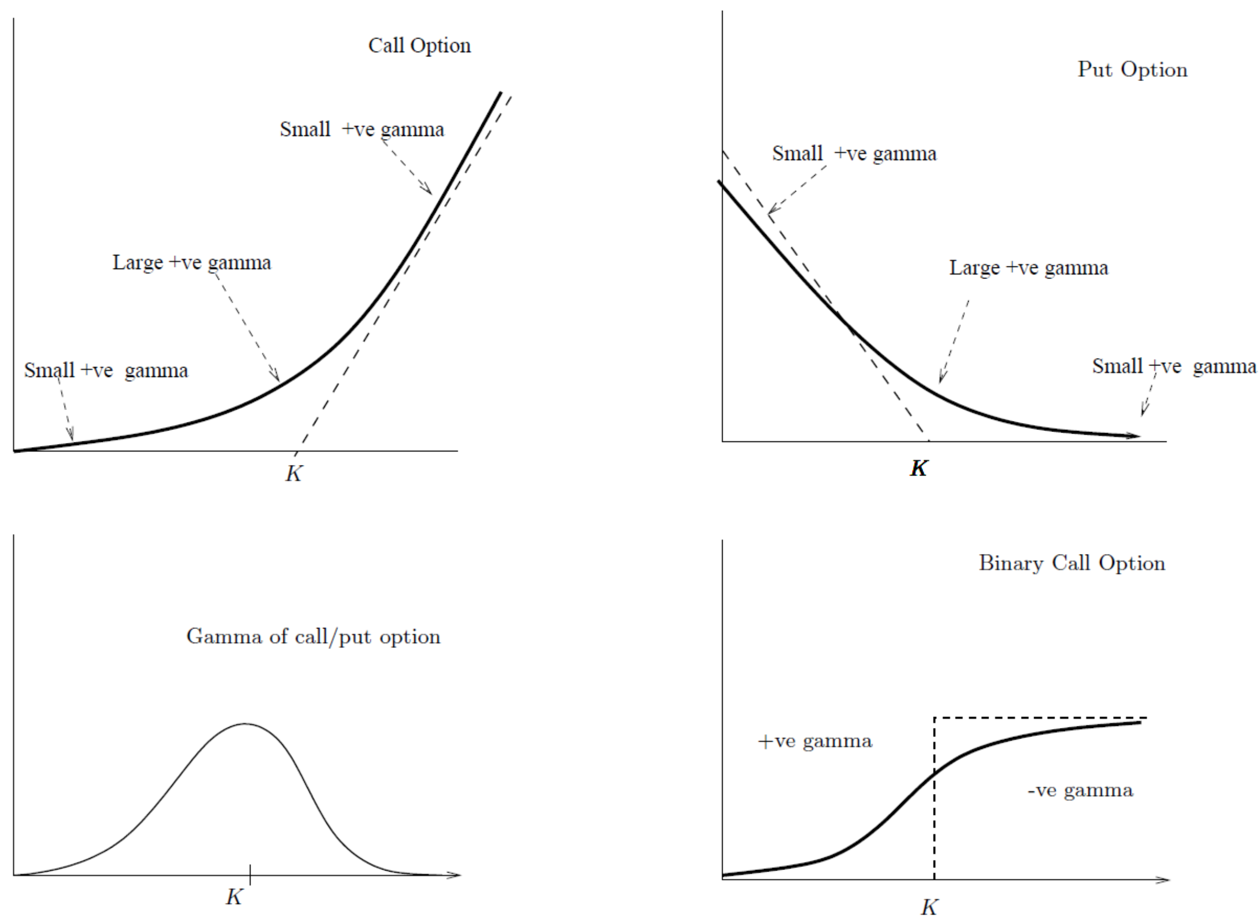

Figure 8.5: Gamma of various options

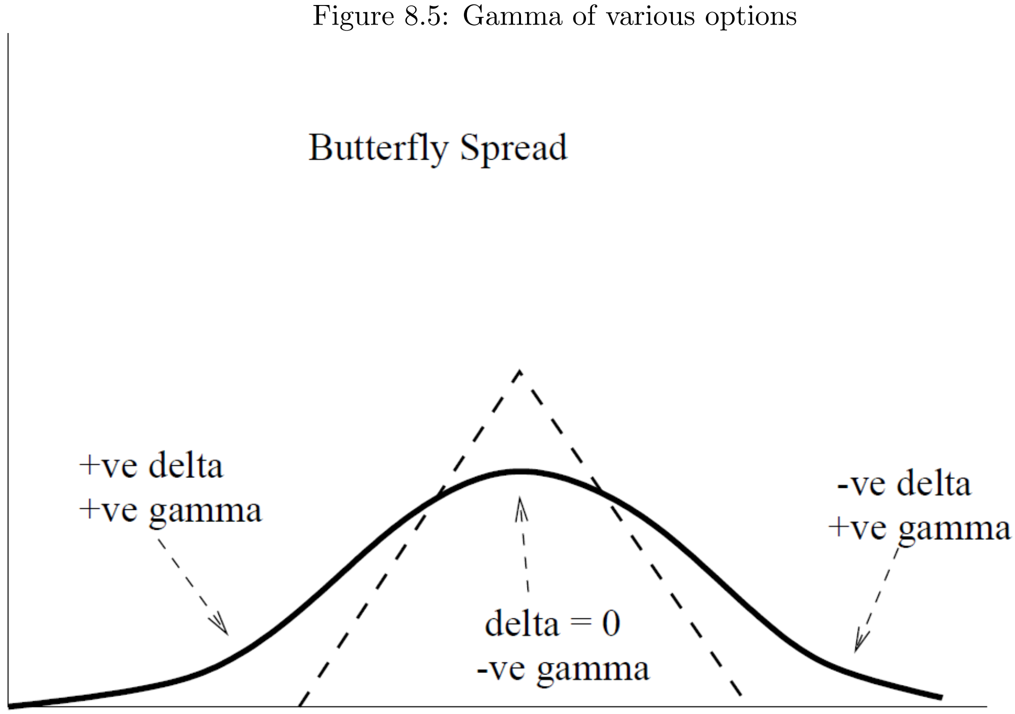

Figure 8.6: Gamma and Delta of a butterfly spread

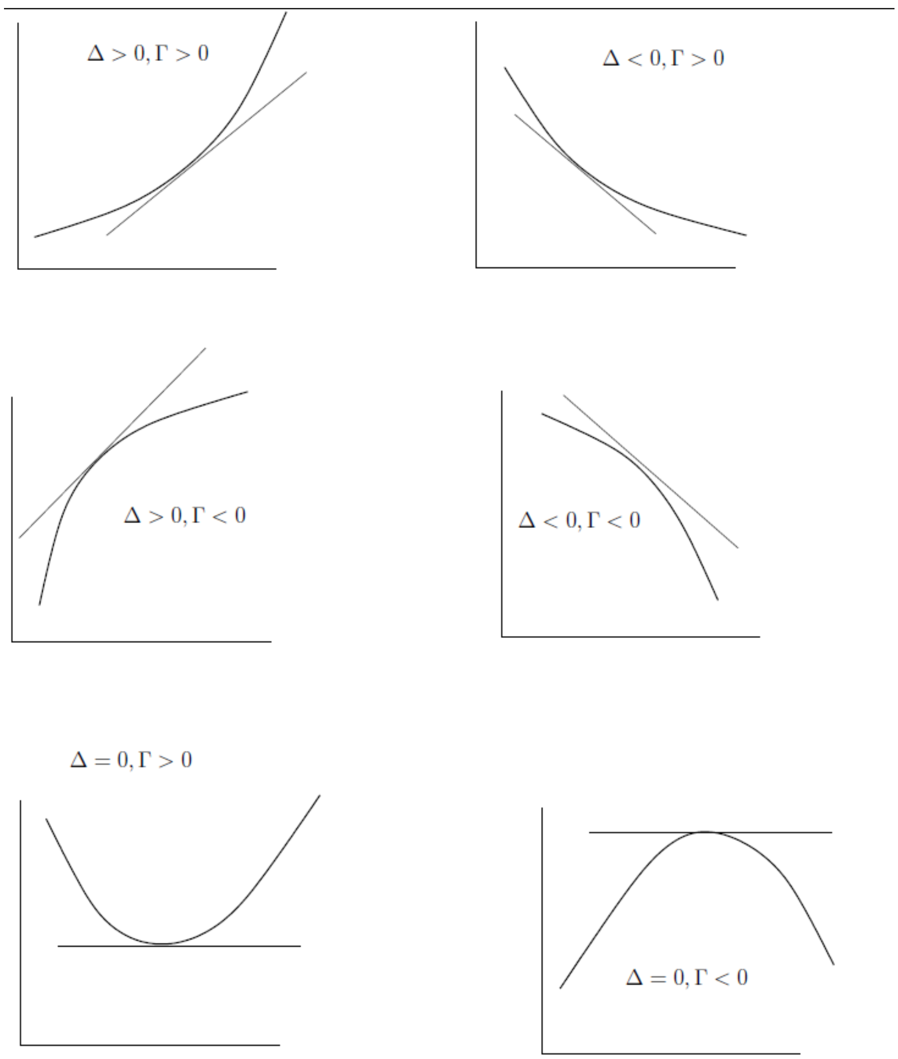

Figure 8.7: The response of the value of a derivative to changes in the underlying asset for different combinations of Δ and Γ.

the derivative loses less value than it would if the price function was uncurved, whereas if S decreases, the derivative gains more value than it would otherwise. Thus whether Δ is positive or negative, positive Γ has a beneficial effect on the change in value of a derivative as a function of the underlying asset.

A similar argument applies to the other cases: If Γ is negative, then a derivative will gain less value than it would if the price function is uncurved, and lose more value. Thus negative Γ adversely affects the change in value of a derivative as a function of the underlying asset.

Note that if a derivative (or portfolio) is delta–neutral, then both a small increase and a small decreas in the underlying asset result in an increase in the value of the derivative (or portfolio).

Gamma for various securities

Clearly the Γ of the underlying share (with respect to itself) is $\frac{\partial^2 S}{\partial S^2} = 0$. The Γ of a forward contract is zero as well, as you can easily verify.

Next, we consider the Γ of a vanilla options. Let C, P be, respectively, a call and a put on an underlying share S (with dividend yield q). From Figure 8.6, it is clear that the gamma of both (long) calls and puts is always positive. Moreover, the call- and put prices are most highly curved around the strike price (at–the–money), and much less curved deep-in and deep–out–of–the–money. Thus the Γ of vanillas options is biggest when the option is near–the–money, and is close to zero for deep–in and deep–out–of–the–money options.

It is left as an exercise to show that

$$\Gamma_C = \frac{N'(d_+)e^{-qT}}{S_0\sigma\sqrt{T}} = \Gamma_P$$

where

$$d_+ = \frac{\ln\frac{S}{K} + (r - q + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

First derive the result for a call, and then use put–call parity to derive the result for a put.

The above formulas for Γ also hold for options on assets which behave like a dividend–paying asset: For foreign currencies, put $q = r_f$, the foreign riskless rate, and for futures put $q = r$.

Note that $N'(x)$ is just the density function of a standard normal random variable. It should therefore come as no surprise that the graph of the Γ of vanilla options as a function of S looks like a bell curve. This is indeed observed in Figure 8.3. Note that Γ is a maximum around the strike price, and that the curve becomes more peaked as the time to maturity decreases. Thus at–the–money options with short maturity have large Γ. Intuitively, this makes a lot of sense: If a trader is delta hedginga vanilla call option, he would like to be long a share if the call expires in–the–money, and to own no shares if the call expires out–of–the–money. The greatest uncertainty in the number of shares that should be owned occurs when the share is at–the–money, where $\Delta \approx 0.5$. Therefore, if the maturity of the option is but a short while away, then an increase in the share price will cause the trader to buy nearly 0.5 extra shares, whereas a decrease will cause him to sell nearly 0.5 shares. Thus to hedge effectively, the rebalancing frequency will be greatest when the option is at–the–money with short maturity.

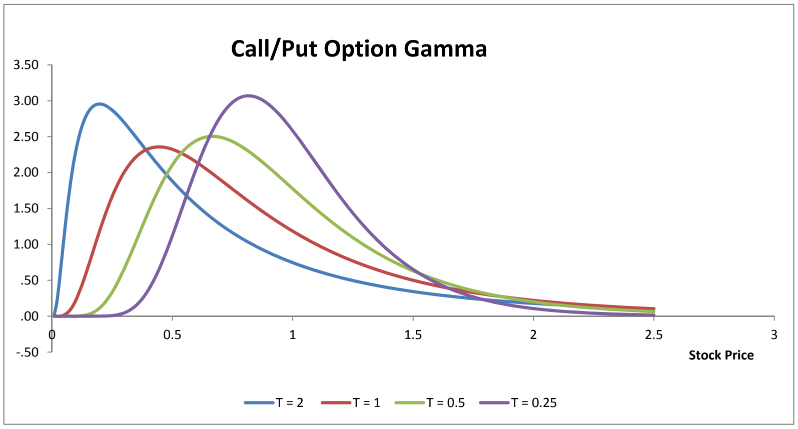

Figure 8.8: Gamma of vanilla options with strike 1 as a function of S, with various maturities.

Hedging with Gamma

Supose that a trader wants to hedge against the value of a portfolio of shares and derivatives losing value due to changes in the underlying asset. To do so, he will need to set the delta of the portfolio equal to zero, making the portfolio delta–neutral. This can be accomplished, for example, by buying or shorting just the right amount of the underlying asset: If the Δ of the porfolio is 23, then shorting 23 shares will make the portfolio delta–neutral.

As time goes by, however, the delta of the portfolio will change in response to changes in the underlying asset, and the trader will need to keep rebalancing his portfolio to maintain a delta–neutral position. If the gamma of the portfolio is very big (positive or negative), then the delta will change very rapidly in response to moves in the underlying asset. If the portfolio has a large gamma, it will therefore need to be rebalanced very frequently, thus incurring large transaction costs.

A trader may therefore seek to make the portfolio simultaneously delta– and gamma neutral. This, however, cannot be done by buying or selling the underlying share: The Γ of the underlying asset is zero, so adding or subtracting shares will not change the gamma of the portfolio. Instead, the trader will need to use a non–linear instrument, such as an option.

Example 8.3.1 A trader has a portfolio Π consisting of various exotic derivatives on a share S. The portfolio has made a tidy profit, and she decides to lock in this profit by hedging away the risk in moves of the underlying share. To ensure that she does not have to rebalance the portfolio too often, she will add shares and vanilla options to the portfolio to make it delta–gamma neutral.

Suppose that

$$\Delta_V = 23 \qquad \Gamma_V = -5$$

The trader decides to first make the portfolio gamma–neutral, by using a call option C which has

$$\Delta_C = 0.5 \qquad \Gamma_V = 1.2$$

By buying 4 call options, the portfolio Γ is made zero. However, each call adds some Δ to the portfolio: The portfolio now has $\Delta_\Pi = 23 + 4 \times 0.5 = 25$. By shorting 25 shares, she can make the portfolio delta–neutral. Moreover, since each share has zero gamma, the gamma of the new portfolio remains zero as well.

Note that the order of hedging is important here: If the trader had first made the portfolio delta–neutral by shorting 23 shares, the addition of 4 call options would have made the portfolio delta non–zero once more.

□

8.4 Theta

The Θ of a derivative or portfolio V measures its sensitivity to changes in time:

$$\Theta = \frac{\partial V}{\partial t}$$

Θ is often called the *time–decay* of a derivative. It is generally negative for options — options are a kind of insurance, and, intuitively, insurance should be cheaper for shorter periods of time. Thus as an option approaches maturity, it should tend to become cheaper (assuming, of course, that there are no changes in the price of the underlying asset). Another way of saying that the option is becoming cheaper is to say that it is *losing value*. Recall that the value of an option can be decomposed into two terms, intrinsic value and time value. Time value tends to decrease as the maturity of the option approaches, thus causing the value of the option to diminish.

Thus Θ is generally negative for options, although deep–in–the–money puts are an exception — se Figure 8.4. Of course, if Θ is negative for a long call, then Θ is positive for a short call. Thus portfolios of derivatives may have positive or negative Θ.

Long American options *always* have negative Θ. This is true for exotic American options as well as vanillas. Since the owner of a two–year American barrier option has the same rights as the owner of a one–year American barier option, and more, the value of a the two–year option *must* exceed that of the one–year option. The same is true for American vanilla options, American binary options, American lookback options, etc. Hence, all other things being equal, the value of an American option decreases as the maturity of the option approaches, so that Θ is negative.

Since an American call and a European call (on a non–dividend paying asset) have the same value, they have the same Θ. It follows that the Θ of European calls must be negative. With put options, things are slightly different. American put options always have negative Θ, but deep–in–the–money European puts may have positive Θ.

The Θ of an in–the–money call option on an asset with a continuous dividend–yield may be positive, however. This is particularly the case if the dividend yield is much greater than the riskless rate. Intuitively, this can be seen as follows: To show that the call has positive Θ, we must show that the value of a call with a long maturity has less value than an otherwise identical call with shorter maturity. Now the call price is just the cost of (continuously) delta

hedging it (ignoring transaction costs). If the call is in–the–money, a delta–hedger will hold a (largish) positive number of shares. Those shares will provide positive income (in the form of the dividend yield). If the dividend exceeds the cost of borrowing, it may actually be cheaper to hedge a long term option than a short term option. Thus the long term option may be cheaper, provided that the underlying provides a positive cashflow, and the hedger holds a largish positive position in the underlying.

Similarly, an in–the-money call on a foreign currency which pays a high interest rate may have positive Θ.

Would a similar argument not apply to American call options on an asset paying a large continuous dividend? No! If an option has positive Θ, then it will gain in value as time approaches maturity (keeping the other variables constant). Thus, where Θ is positive, the graph of the option price must lie below the graph of the payoff. Now the payoff function is just the *intrinsic value* of the option, and the value of an American option can never be less than its intrinsic value. Thus American options cannot have positive Θ. If Θ is positive, the option is worth less than its intrinsic value, and should be exercised immediately.

Theta for various securities

Note that the Θ of the underlying share is zero. The calculation of the Θ of a forward contract is straightforward: Since the time–t value of a forward contract with delivery date T and delivery price K is just $V = S_t - Ke^{-r(T-t)}$, we see that

$$\Theta = \frac{\partial V}{\partial t} = rKe^{-r(T-t)}$$

It is left as an exercise to show that the Θ's of a vanilla European call C and a European vanilla put P on an asset S with continuous dividend yield q are given by:

$$\begin{aligned}
\Theta_C &= -\frac{S_0 e^{-qT} N'(d_+)\sigma}{2\sqrt{T}} + qS_0 e^{-qT} N(d_+) - rKe^{-rT} N(d_-) \\
\Theta_P &= -\frac{S_0 e^{-qT} N'(d_+)\sigma}{2\sqrt{T}} - qS_0 e^{-qT} N(-d_+) + rKe^{-rT} N(-d_-)
\end{aligned}$$

where d_+, d_- are as usual. Observe that

$$\Theta_P = \Theta_C + rKe^{-rT} - qS_0 e^{-qT}$$

The above formulas for Θ also hold for options on assets which behave like a dividend–paying asset: For foreign currencies, put $q = r_f$, the foreign riskless rate, and for futures put $q = r$.

Hedging with Theta

In order to protect a derivative or portfolio from the risk that the underlying share price will change, a trader will delta hedge. One cannot exactly call time a *risk–factor*, however, since there is no uncertinty as to what the time will be tomorrow or next month. Hedging against the passing of time thus makes little sense. Nevertheless, Θ is an important descriptor of the behaviour of a derivative or portfolio over time. Moreover, if a portfolio Π is rebalanced to maintain a delta–neutral position, there is a strong relationship between Θ and Γ — If $\Delta = 0$, the Black–Scholes PDE becomes

$$\Theta + \frac{1}{2}\sigma^2 S^2 \Gamma = r\Pi$$

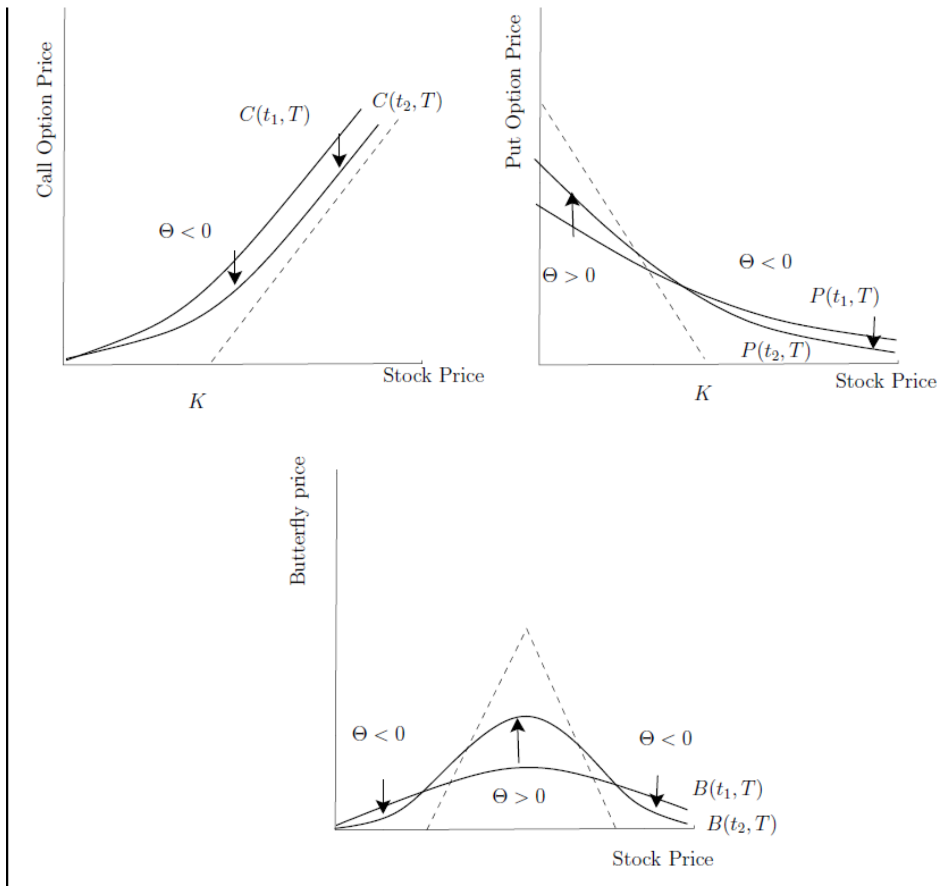

Figure 8.9: The behaviour of call, put, and butterfly spread prices as time approaches maturity. Here $t_1 < t_2 < T$. It is clear from the diagram where Θ is positive, and where it is negative.

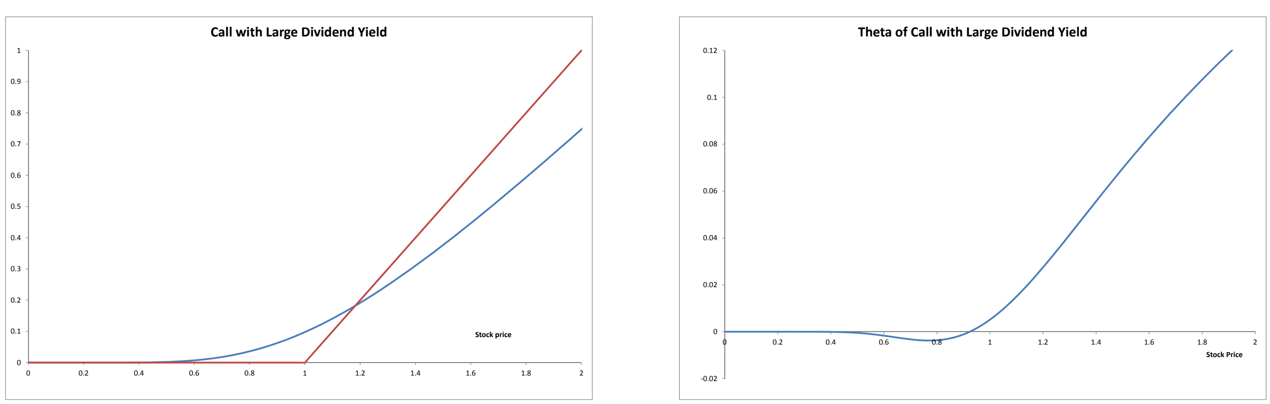

Figure 8.10: Theta of a call on an asset that has a large dividend yield.

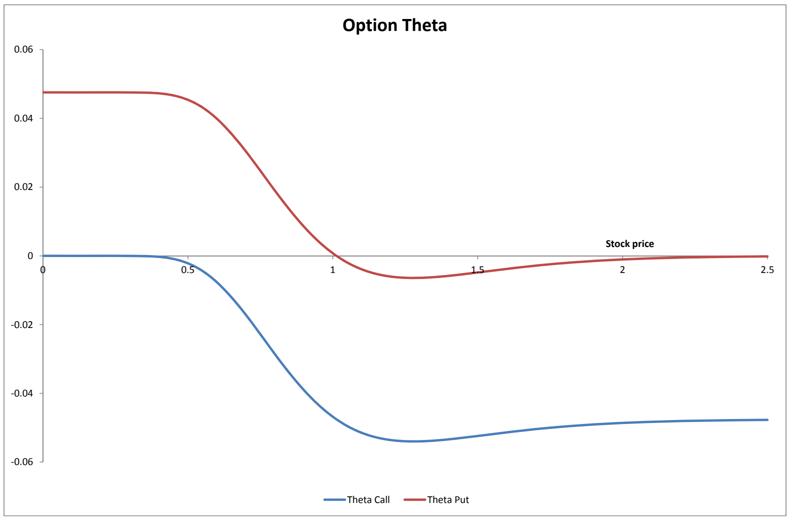

Figure 8.11: Theta of calls and puts as a function of the underlying asset.

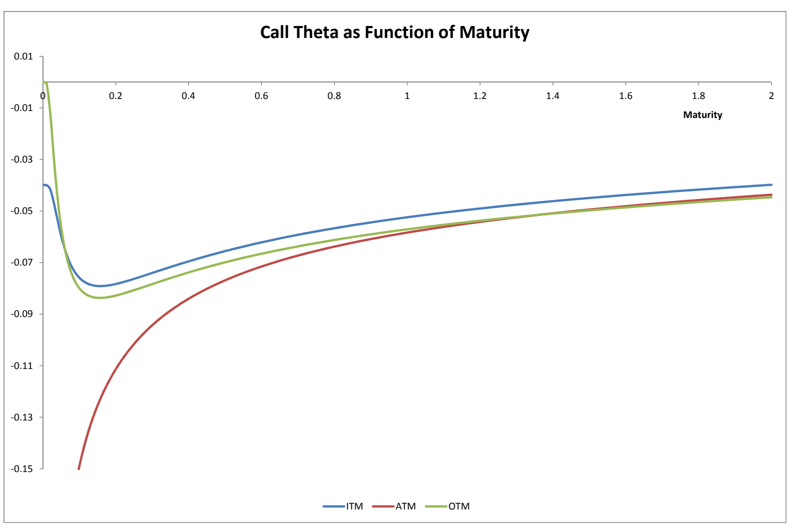

Figure 8.12: Theta of calls and puts as a function of the underlying asset.

If Π is delta–hedged, its value will remain approximately constant over time. Thus an increase in Θ will imply a decrease in Γ, and vice versa.

8.5 Vega and Rho

In the derivation of the Black–Scholes formulas, we assumed that the volatility and the riskless interest rate are constant. In reality, this is not so. We have already explained why options on an asset with high volatility should be more expensive than options on an asset with low volatility (assuming identical strikes, maturities and initial asset prices): Options are a form of insurance, and protection against a high risk event costs more than protection against a low risk event. The volatility of an asset is a measure of how risky an investment it is. We also know that an increase in interest rates tends to cause an increase in call prices and a decrease in put prices. If rates increase, $PV(K)$ becomes smaller, so that a party who is long a call or short a put benefits.

Nevertheless, neither volatility nor the riskless rate are constant in the real world — it is here where the Black–Scholes model breaks down, and there exist many models which try to capture changing volatilities and/or rates. Presently, no one model seems superior to the rest, and traders still work in the Black–Scholes framework. Instead of using the *historical volatility*[2] of an asset, they will use their own view of what the volatility of the underlying asset will be like over the life of the option. Options are therefore often *quoted* in volatility, not price. If a trader says that he is prepared to sell a call at 35% volatility, this gives more immediate information than if he says that he is prepared to sell it for \$1.23. In the same way that yield is a better measure to compare bonds than price, volatility is a better measure to compare options than price.

It is therefore important to know how the value of a derivative V will be affected by changes in volatility. Vega, denoted $\mathcal{V}$ measures just that:

$$\mathcal{V} = \frac{\partial V}{\partial \sigma}$$

It is left as an exercise to show that the $\mathcal{V}$'s of a European call C and a European put P on an asset S with a dividend yield q are given by:

$$\mathcal{V}_C = S_0 e^{-qT} \sqrt{T} N'(d_+) = \mathcal{V}_P$$

(That calls and puts have the same $\mathcal{V}$ follows directly from put–call parity.) The above formulas for $\mathcal{V}$ also hold for options on assets which behave like a dividend–paying asset: For foreign currencies, put $q = r_f$, the foreign riskless rate, and for futures put $q = r$.

Note that $\mathcal{V}$ is positive for long vanilla European options. This proves that, in the Black–Scholes world, an increase in volatility implies an increase in option value. Since $N'(x)$ is just the bell–shaped density function of a standard normal random variable, the graph of $\mathcal{V}$ is bell–shaped as well. $\mathcal{V}$ diminishes as time approaches maturity, and is greatest at–the–money.

Indeed, for vanilla options Γ, Θ and $\mathcal{V}$ are all of maximum size for near–the–money options. Thus it is only for near–the-money options that these hedging sensitivities become important.

Finally,

$$\rho = \frac{\partial V}{\partial r}$$

[2] i.e. the volatility gleaned from historical market prices

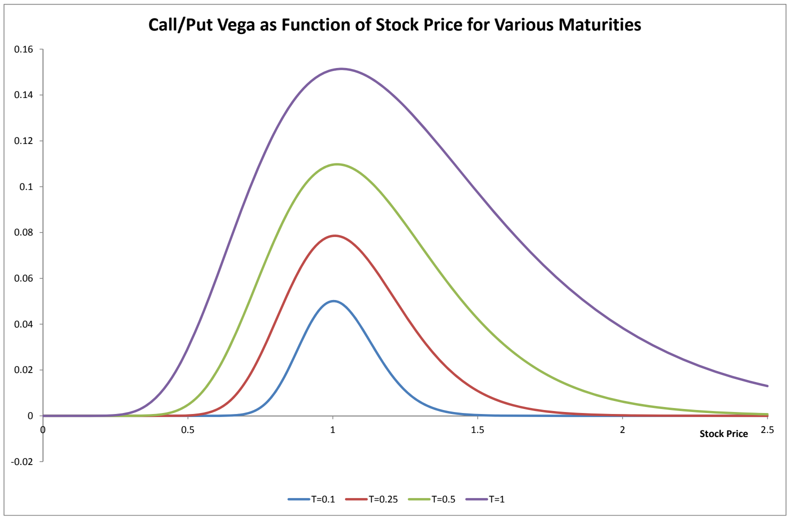

Figure 8.13: Vega for European vanilla options of various maturities.

measures the sensitivity of a derivative V to changes in the interest rate. It is not hard to show that for calls C and puts P on an asset S with dividend yield q we have

$$\rho_C = KTe^{-rT}N(d_-) \qquad \rho_P = -KTe^{-rT}N(-d_-)$$

The graph of ρ looks a bit like the distribution function of a normal random variable $N(x)$. Note that ρ is always positive for European call options, and always negative for European put options. This proves that, in the Black–Scholes model, an increase in interest rates causes an increase in the value of a call and a decrease in the value of a put.

8.6 Implied Volatility and the Volatility Skew

The Black–Scholes option pricing formulas require as input parameters S_0, K, T, r and σ. The volatility σ is the standard deviation of annualised returns of the stock, and is the only parameter that cannot be directly observed, but must be gleaned from market data. For example, one can calculate the daily returns $\frac{S_{t+1}-S_t}{S_t}$ of a share S for a year, and then find the standard deviation of daily returns. The annualised return can then be obtained from the daily return by multiplying by $\sqrt{365}$, or better yet, by $\sqrt{\text{no. of trading days p.a.}}$

The main problem with this approach is that it calculates the volatility of the share as it was in the past: *historical volatility*. But option traders do not care about past volatility — it is future volatility, volatility over the life of the option, that is important. In practice, therefore, traders need to include their view about the future volatility of the share into the option price.

Given the *market* price of a call option C, one can ask the following question: For which value of σ will the *theoretical* Black–Scholes price of the option be equal to the market price?

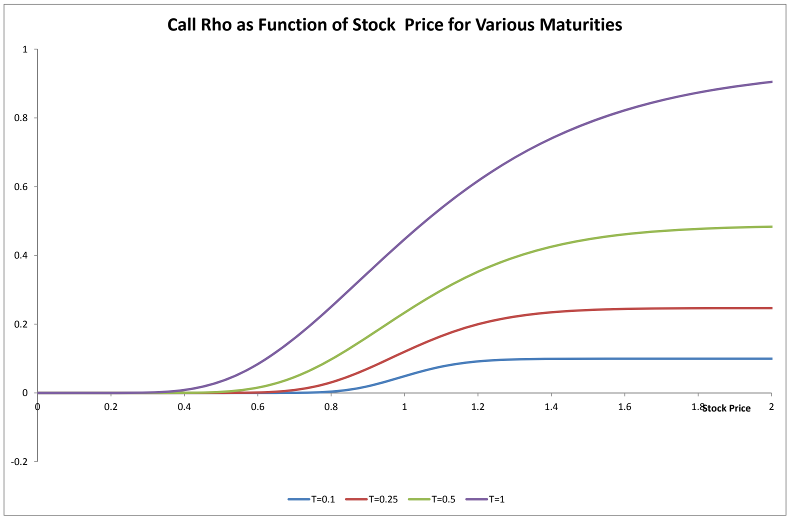

Figure 8.14: Rho for European call options of various maturities.

This is the volatility *implied* by the market price of the option, and is therefore called the *implied volatility* of the asset. There is no–closed form solution of implied volatility; instead, a numerical procedure such as Newton's method (or Excel's SOLVER) must be used to calculate the implied volatility of an option given its price.

Many exchange–traded options are extremely liquid, and their prices therefore represent some kind of market consensus view about volatility: The implied volatility measures how risky a large proportion of traders think the market will be in the future. Thus implied volatility has a *term structure*: If the implied volatility of a one–year call is lower than that of a 6–month call (both with the same strike), it is because traders think that the "average" volatility over one year will be less than the "average" volatility over 6 months.

Moreover, if one calculates the volatility implied by a one–year call with strike 100 and the volatility implied by a one–year call with strike 110, one may not get the same answer! There is a so–called *volatility smile* or *volatility skew*, as shown by Figure 8.6.

Does this mean that traders who trade options with strike 100 have a different view about future market volatility than traders who trade in options with strike 110? No. The problem here is that traders do not believe that the Black–Scholes model is correct. The Black–Scholes model assumes that the underlying asset price follows a geometric Brownian motion, and thus that future asset prices are lognormally distributed. Traders do not believe that this is the case. For example, they may believe that very large moves in the asset price are more likely than those predicted by the Black–Scholes model. They may also believe that very small moves are more likely. Of course, together these mean that traders believe that average–sized moves are less likely than those predicted by the Black–Scholes model. The implied distribution of asset prices therefore has a higher peak near the current asset price, and has heavier tails — statisticians call such a distribution *leptokurtic*. This is the situation shown on the left in Figure 8.6.

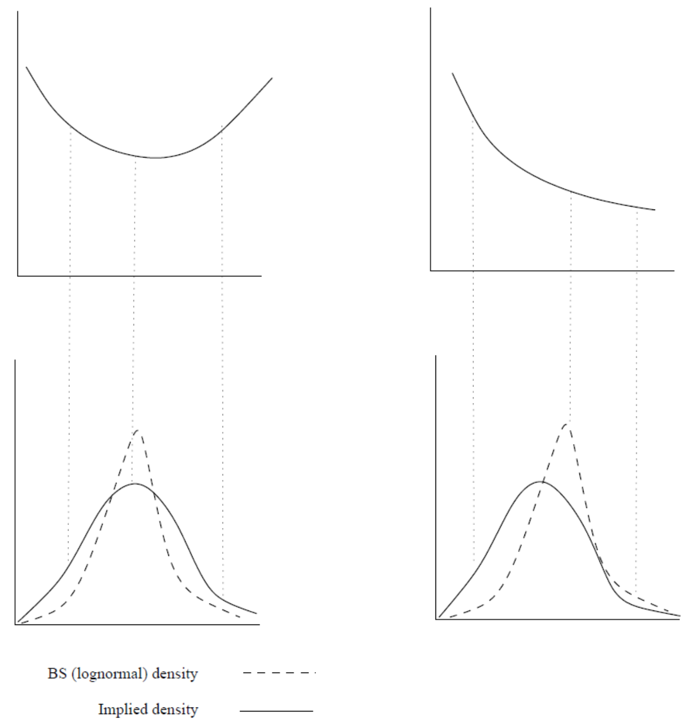

Figure 8.15: The volatility smile

Before we show how the volatility skew is explained by a different implied distribution, just note the following: The put–call parity formula for European vanilla options was derived long before we modelled stock prices as geomeric Brownian motions. Thus whereas the Black–Scholes pricing formula is correct only if our stock price dynamics are correct, the put–call parity is independent of atock price models: If put–call parity fails to hold, ther is arbitrage, no matter what anyone thinks might be the “correct” model of stock prices.

Now the Black–Scholes prices of calls and puts (with the same strike and maturity) satisfy the put–call parity relations. Thus if we calculate the implied volatility using a call with strike K an maturity T, and then calculate the implied volatility using a one year put with the same strike and maturity, we should get the same answer. If we calculate different implied volatilities, put–call parity would fail, and there would be arbitrage. Therefore the graph of vaolatility vs strike for calls should be the same as the graph of volatility vs strike for puts.

If the right tail of the implied asset distribution is fat, then a call with a high strike price (deep–out–of–the–money)is more likely to expire in–the–money than predicted by the Black–Scholes model. It should therefore cost more than the theoretical Black–Scholes price. Now since option price is an increasing function of volatility, a more expensive option has a higher volatility.

Similarly, a deep–out–of–the–put call is more likely to expire in–the–money if the left tail of the distribution is fat. It should therefore be more expensive than a Black–Scholes world option. Finally, a near–the–money call or put option is less likely to expire in–the–money than what the Black–Scholes model predicts. Near–the–money options are therefore cheaper than Black–Scholes world options.

Remarks 8.6.1 You may suspect that all is not well with the above argument. True, a deep–out–of–the–money put is more likely to expire in–the–money if the left tail is fat. But isn't a deep–in–the–money call then more likely to expire out–of–the–money?

True.

Shouldn't a deep–in–the–money then be cheaper than what is predicted by the Black–Scholes model, so that the implied volatility is lower?

No. Firstly, this would mean that the volatility implied by a deep–out–of–the–money put and a deep–in–the–money call would be different: The put price is higher, so implies a high volatility, and the call price would be lower[3], so implies a low volatility. But calls an puts with the same strike and maturity must have the same implied volatility, or else there is arbitrage.[4].

Secondly, remember that the option price is determined by two factors:

(1) The probability that the option expires in–the–money, and

(2) The expected payoff given that the option expires in–the–money.

For example, we argued that an increase in volatility leads to an increase in option price as follows: If the volatility increases, then the option has a greater chance of expiring deep in–the–money. Of course it also has a greater chance of expiring deep–out–of–the–money (and thus a smaller chance of expiring near–the–money). But the probability of expiring deep–out–of–the–money does not affect the option price. The price of a call is

$$C_0 = e^{-rT}\mathbb{E}[(S_T - K)^+] = e^{-rT}\left[\mathbb{E}[S_T; S_T \geq K] - K\mathbb{P}(S_T \geq K)\right]$$

where the expectation and probability are *risk–neutral*. So the option price is is determined solely by the probability that it expires in–the–money, and the value of the underlying given that the option expires in–the–money.

Hence we should look an option's behaviour when it expires in–the–money to find the price. The fact that a call is more likely to expire out–of–the–money when the left tail is fat is therefore *irrelevant* to option pricing.

□

[3] If this argument is correct…

[4] …so this argument is not correct

Cumulative Normal Distribution

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.543795	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.659097	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.8790	0.8810	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.9032	0.9049	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.9222	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.9452	0.9463	0.94738	0.94845	0.9495	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.9608	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.9732	0.97381	0.97441	0.9750	0.97558	0.97615	0.9767
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.9803	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.9830	0.98341	0.98382	0.98422	0.98461	0.9850	0.98537	0.98574
2.2	0.9861	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.9884	0.9887	0.98899
2.3	0.98928	0.98956	0.98983	0.9901	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.9918	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.9943	0.99446	0.99461	0.99477	0.99492	0.99506	0.9952
2.6	0.99534	0.99547	0.9956	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.9972	0.99728	0.99736
2.8	0.99744	0.99752	0.9976	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.9990
3.1	0.99903	0.99906	0.9991	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.9994	0.99942	0.99944	0.99946	0.99948	0.9995
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.9996	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.9997	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.9998	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.9999	0.9999	0.9999	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997
4.0	0.99997	0.99997	0.99997	0.99997	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998

Table for $N(x)$ when $x \geq 0$.

Notes:

(1) Note that $N(-x) = 1 - N(x)$.

(2) Use linear interpolation to calculate the values of $N(x)$: For example,

$$N(0.8625) = N(0.86) + 0.25(0.87 - 0.86) \approx N(0.86) + 0.25[N(0.87) - N(0.86)]$$